USING EARNED VALUE DATA TO DETECT POTENTIAL PROBLEMS IN ACQUISITON CONTRACTS

THESIS

C. Grant Keaton, Second Lieutenant, USAF

AFIT/GCA/ENC/11-01

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

AFIT/GCA/ENC/11-01

USING EARNED VALUE DATA TO DETECT POTENTIAL PROBLEMS
IN ACQUISITION CONTRACTS

THESIS

Presented to the Faculty

Department of Mathematics and Statistics

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degree of Master of Science in Cost Analysis

C. Grant Keaton

Second Lieutenant, USAF

March 2011

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

AFIT/GCA/ENC/11-01

USING EARNED VALUE DATA TO DETECT POTENTIAL PROBLEMS
IN ACQUISITION CONTRACTS

C. Grant Keaton
Second Lieutenant, USAF

Approved:

________//signed//_________________ 28 Feb 2011
Edward D. White, PhD (Chairman) Date

________//signed//_________________ 1 Mar 2011
Lt Col Eric J. Unger, PhD (Member) Date

________//signed//_________________ 28 Feb 2011
Gilbert L. Peterson, PhD (Member) Date

Abstract

Government contractors report earned value (EV) information to government agencies in monthly Contract Performance Reports (CPR). Though major differences may exist in the data between subsequent CPRs, we know of no government effort to detect these occurrences. The identification of major changes may locate and isolate problems and thus prevent million and billion dollar cost and schedule overruns. In this study, we develop an approach to identify changes in the Cost Performance Index (CPI) and the Schedule Performance Index (SPI) that may indicate problems with contract performance. We find the detection algorithm indentifies changes in the CPI and the SPI that correspond to large future changes in the Estimate at Complete (EAC). The ability to detect unusual changes provides decision-makers with warnings for potential problems for acquisition contracts.

Acknowledgements

My thesis is the result of hard work, although not all my own. My committee gave me an opportunity to benefit from their knowledge and experience. Thank you Dr. White, Lt Col Unger, and Dr. Peterson.

Members of ASC, AFIT, DCARC, and other organizations gave me the gift of their time. Specifically, thank you Col. Jim Reitzel, Mr. Randal Bowen, Mr. Jim Davis, Mr. Stan Nisevich, Dr. Ross Jackson, Dr. Joseph Pignatiello, Mr. Jordan Goldmeier, Mr. Alex Gutman, Mr. Burt LeClercq, and Mr. John McGahan.

Thank you Mr. Steve Miller. Your sponsorship of my thesis research provided the ability to mature an idea.

My classmates and program director guided me through my first year and a half as an Air Force officer. They helped me realize important opportunities, navigate some difficult situations, and get past others.

Finally, I want to thank my family, which has always given me immeasurable support. Although their influence on me and my work is less concrete, the nature of their support makes it the most meaningful.

Grant Keaton

Table of Contents

Page

List of Figures

Page

List of Tables

List of Equations

Page

USING EARNED VALUE DATA TO DETECT POTENTIAL PROBLEMS IN ACQUISITION CONTRACTS

1: Introduction

Strains on the discretionary budget force military services to monitor cost and schedule performance for materiel acquisition closely. However, the deterioration of skills and personnel in the defense acquisition workforce decreased the Department of Defense's (DoD) ability to provide adequate financial discipline (Morin, 2010). While DoD is presently addressing the reconstitution of the defense acquisition workforce (Morin, 2010), current acquisition analysts continue to manage an increasing workload. These analysts require new approaches to improve financial discipline in defense acquisition.

Several methods exist that may improve acquisition analysts' ability to monitor cost and schedule performance. Specifically, analysts may develop more accurate Estimate at Complete (EAC) models and scrutinize changes in cost and schedule performance indices (Christensen, Antolini, & McKinney, 1995). Improvements in these methods continue research on the results of poor cost and schedule performance, not the identification of symptoms one requires for real-time correction. If analysts can identify potential and actual problems instead of their symptoms, program managers can monitor high-risk activities diligently to prevent poor cost and schedule performance.

Our Contribution

This research provides program analysts and DoD leadership with an approach for identifying problems within acquisition contracts in real-time. At a high-level, we discard the typical approach to acquisition research by treating earned value data as a general data time series, not as program performance measures with definite interpretations. Specifically, we test the ability of a forecasting algorithm to detect statistically significant changes in acquisition contracts' Cost Performance Index (CPI) and Schedule Performance Index (SPI). Successful models will identify contract areas which are at risk of or face ongoing cost overruns and schedule delays. Although program managers can use this information to aid analysis, this approach is not a substitute for in-depth understandings of their programs.

Particularly, we center our research on the following questions:

1. Can we detect changes in acquisition contracts with a detection algorithm given at least the first three months CPI and SPI data?
2. If we can detect changes, how long does a change exist before we identify it?

In the next chapter, we discuss change detection research, time series forecasting, and analysis of earned value data. Chapter III reviews our methodology in-detail. Particularly, we discuss earned value data, Autoregressive/Integrated/Moving Average (ARIMA) models, and the change detection algorithm. Chapter IV presents the detection results and relationships between changes in the CPI and with SPI with major changes in the Estimate at Complete (EAC). For different algorithm sensitivities, we detect between 10% and 60% of major changes in the EAC that occur in the same month as the

detection. Additionally, we find 20% to 50% of detections correspond to major changes in the EAC in future months. Finally, Chapter V summarizes the significant findings of the research, discusses implications to DoD policies, and suggests areas of future research.

II. Literature Review

Researchers apply change detection to identify when system characteristics change. The wide applicability of the technique makes change detection less an academic field than a methodology many fields use for analysis. Signal processing (Borodkin & Mottl', 1976) (Cohen, 1987), time series analysis (Box, Jenkins, & Reinsel, 1994)(Dasgupta & Forrest, 1996), automatic control (Willsky, 1976), and industrial quality control (Shewhart, 1931) (Woodward & Goldsmith, 1964) (Duncan, 1986) are some fields that apply change detection techniques. However, increases in information availability and advances in computer processing power provide new opportunities for change detection research (Cios & Moore, 2002) (Venkatesh, 2007).

Change detection techniques hinge on the definition of system change. A single definition does not exist because researchers interpret change differently within and across fields. In spite of the various interpretations of change, typically definitions of change detection focus on time-dependency. Specifically, abruptness, not necessarily magnitude, characterizes system change. (Basseville & Nikiforov, 1993).

The design of a change detection system is an important element of the technique. Different detection capabilities require detection system designers to balance the general and the specific applicability of a model. To achieve this balance, system designers accept tradeoffs between certain detection performance characteristics.

Frequently, change detection researchers devise and appraise models with the following intuitive performance indices:

1. Mean delay for detection
2. Mean time between false alarms
3. Probability of non-detection
4. Probability of false alarms
5. Accuracy of change time and change magnitude estimates (Basseville & Nikiforov, 1993)

Another consideration in change detection is the type of problem a system attempts to solve. An online approach focuses on real-time solutions because the model treats information serially. Consequently, the online approach can identify non-optimal solutions because the approach does not use an entire input data stream and thus searches for local optimality (Borodin & El-Yaniv, 1998) (Gustafsson, 2000). Often researchers who use online change detection algorithms use performance criteria based on the mean delay for detection and the mean time between false alarms (Basseville & Nikiforov, 1993). These performance criteria adjust the detection capability of the algorithm toward instantaneous, though sometimes incorrect, identification of change.

Alternatively, offline models offer retrospective analysis of changes in system characteristics. This approach requires complete input data streams to search for globally optimal solutions (Gustafsson, 2000). Researchers further divide offline detection into the evaluation of change-no change hypotheses tests and the estimation of change time. Change-no change hypotheses tests attempt to maximize the probability of correct change detection with a certain probability of incorrect change detection. Change time estimation determines the maximum probability the actual change time occurs within a definite confidence interval (Basseville & Nikiforov, 1993).

Time Series Analysis

Time series analysis offers an approach to both online and offline change detection (Makridakis, Wheelwright, & Hyndman, 1998). Equally important, time series analysis addresses dependency often found in observations at distinct intervals of a time series. The combination of these analysis capabilities allows researchers to study common time-dependent problems with the technique. Specifically, researchers address four practical problems with time series analysis:

1. Forecast future values using past and current observations
2. Monitor the effect dynamic inputs have on an output
3. Examine how disturbances to input variables effect the behavior of a time series
4. Adjust input variables to compensate for output deviations (Box, Jenkins, & Reinsel, 1994)

Forecasting

Quantitative forecasting allows researchers to predict future outcomes probabilistically (Makridakis, Wheelwright, & Hyndman, 1998). Implicitly, the value of quantitative forecasting depends on the satisfaction of the assumptions that sufficient lead time exists and known, conditional factors affect the outcome of a final event. Forecasting provides little benefit if lead time or planning does not impact the final outcome or the factors that do affect the final outcome are unknown. Explicitly, quantitative forecasting requires 1) quantifiable information about past events and 2) the expectation at least some earlier patterns will repeat in the future.

Numerous methods of forecasting exist, ranging from the makeshift to the mathematically formal. However, all forecasting models follow the general model of Equation 2.1.

$$observation = pattern + error \tag{2.1}$$

The essential responsibility of any forecaster is to separate the pattern from the error. The successful separation of the two components provides a forecaster with the appropriate pattern to characterize the time series (Makridakis, Wheelwright, & Hyndman, 1998).

Regression sides with formal mathematical forecasting and is one of the most common forecasting techniques. Regression relies on input or explanatory variables to model changes in the outcome or response variable. An important subset of regression is autoregression. With autoregression, one substitutes explanatory variables X_i with earlier values of the forecast variable Y_{t-l}. Equations 2.2 and 2.3 are general form equations for regression and autoregression models, respectively:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots \beta_l X_l + e \tag{2.2}$$

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \cdots \beta_l Y_{t-l} + e_l \tag{2.3}$$

where i denotes a particular explanatory variable, t denotes time, l reflects time lag, β_l is a weighting coefficient, and e is the forecast error. When a time series exhibits relationships between observations of specific intervals, autoregression may be an appropriate technique because it incorporates the relationships and predictive capabilities of prior observations for a present forecast.

Autoregressive/Integrated/Moving Average (ARIMA)

The Autoregressive/Integrated/Moving Average (ARIMA) model is a common forecasting technique which incorporates the autoregressive model with the moving average model and a differencing mechanism. The technique gained prominence during the 1970s when George Box and Gwilym Jenkins published their seminal work *Time Series Analysis: Forecasting and Control*. In their book, Box and Jenkins described the theoretical framework for univariate time series ARIMA models.

Assumptions.

In theory, ARIMA models are the most general class of stationary forecasting models. Despite the broad uses of ARIMA, proper application of this class of models requires strict adherence to the assumptions of ARIMA modeling. Namely, a time series must 1) be stationary in the mean, 2) be stationary in variance, and 3) have a distribution of forecast residuals that is approximately normal with a mean of zero and standard error of $\frac{1}{\sqrt{n}}$, where *n* is the number of observations (Makridakis, Wheelwright, & Hyndman, 1998).

The assumption of a stationary mean and variance in a time series has important implications. The principle concern of stationary time series is that one cannot forecast the characteristics of a non-stationary time series well. For example, if a time series increases over time, the mean and the variance will increase with the number of observations. As a result, forecasts will always underestimate the mean and the variance. Additionally, because the mean and the variance of a non-stationary time series are uncertain, one may infer little about correlations with other variables (Nau, 2005).

To be stationary in the mean, the time series shows no evidence of a change in the mean through time. Similarly, no meaningful changes in the variance over time indicate the variance is stationary-- homoskedasticity. Though violations of these assumptions often are clear visually, the Dickey-Fuller and Augmented Dickey-Fuller unit root tests are robust methods of verification (Makridakis, Wheelwright, & Hyndman, 1998).

Forecasters address violations of the stationary assumptions with difference (or de-trend) and transformation routines. If successful, the forecaster may find a time series is stationary in an alternative view of the data. Typically, an analyst uses difference calculations to adjusts upward or downward trends in the mean of the time series. With a difference or de-trend calculation, the analyst subtracts the previous observation from the current observation to find the difference:

$$\Delta Y = Y_t - Y_{t-1} \tag{2.4}$$

where Y is the observation and t is the time of the observation. Likewise, forecasters use mathematical transformations to address violations of the stationary variance assumption. The type of transformation depends on the specific time series, which include common transformations such as natural logarithms and exponential functions.

Figure 2.1 shows an example of a time series with a non-stationary mean--specifically an uptrend. Figure 2.2 illustrates the effect of a first order non-seasonal difference on the data in Figure 2.1. As a result, the mean is approximately stationary with deviations that tend to revert to the mean. Additionally, the variance of the time series in both Figures 2.1 and 2.2 appears stationary, with no clear indication of a potential to change over time. One can verify these results with unit root tests.

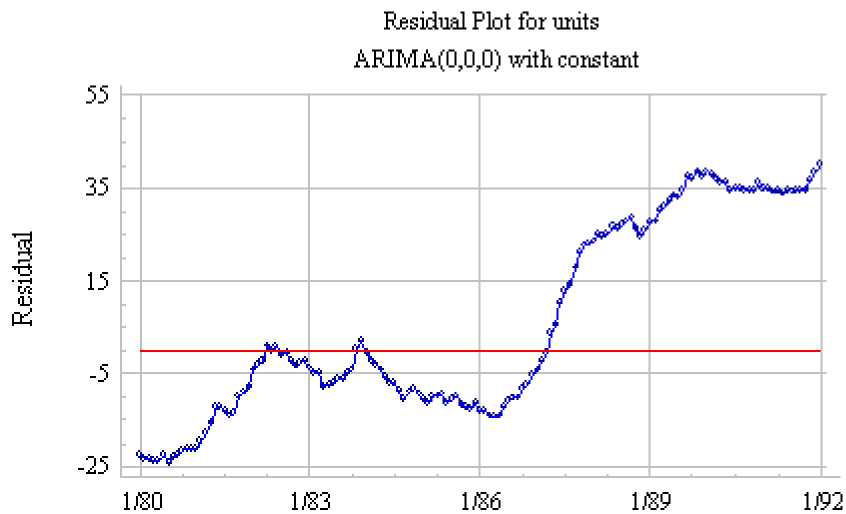


Figure 2.1: Upward Trend Plot (Nau, 2005)

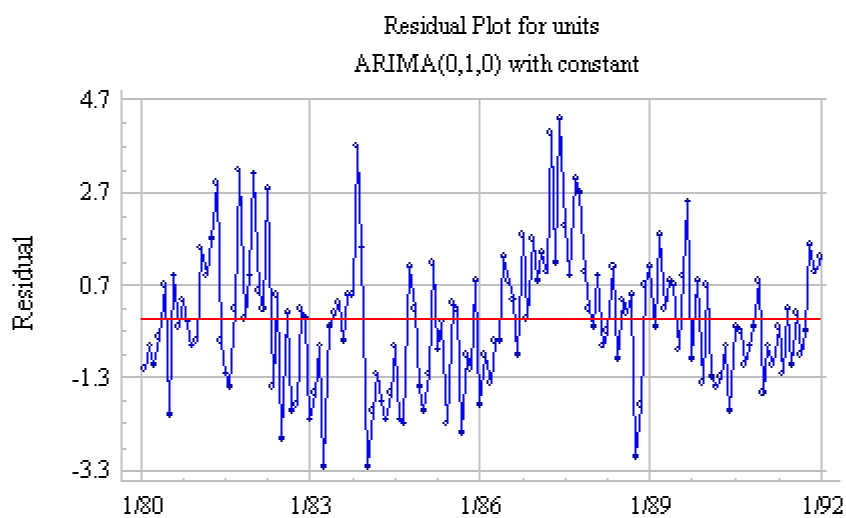


Figure 2.2: First Non-Seasonal Difference (Nau, 2005)

The assumption that the normal distribution approximates the distribution of the forecast residuals is a diagnostic test to ensure the forecast errors truly are random. If this assumption is not met, perhaps the model omits meaningful patterns. Forecasters test normality of the residual distribution with traditional normality and portmanteau tests. One traditional test of normality is the Shapiro-Wilk goodness of fit test. The Shapiro-

Wilk method tests the null hypothesis that a sample $x_1, x_2, \ldots, x_n$ comes from a population with a normal distribution. Equation 2.5 lists the test statistic for the Shapiro-Wilk test:

$$W = {(\sum_{i=1}^{n} c_i x_i)^2}\Big/{\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{2.5}$$

where c_i is a constant, $\bar{x}$ is the sample mean, and x_i is an ith order statistic (Shapiro & Wilk, 1965). Portmanteau test compare the residuals of the autocorrelation function (ACF) and the partial autocorrelation function (PACF) to the normal distribution to ensure the distribution of the residuals is approximately normal. Box and Piece and Ljung and Box developed two common portmanteau tests (Box & Pierce, 1970) (Ljung & Box, 1978). Equations 2.6 and 2.7 list the Box-Pierce and Ljung-Box portmanteau tests, respectively:

$$Q = n \sum_{k=1}^{h} r_k^2 \tag{2.6}$$

$$Q^* = n(n+2) \sum_{k=1}^{h} (n-k)^{-1} r_k^2 \tag{2.7}$$

where n is the number of observations in the time series, h is the number of lag periods the analysts consider, and r_k is the correlation value for observation k. Both portmanteau tests compare Q and Q^* to the chi-square distribution to determine if the plot of the residuals is statistically different from "white noise".

General Non-Seasonal ARIMA Model.

Forecasters describe the non-seasonal ARIMA model as an ARIMA(p, d, q), where:

- p is the number of autoregressive terms,
- d is the number of non-seasonal differences, and
- q is the number of lagged forecast errors (Nau, 2005).

Specifically, autoregressive terms are the lags of a differenced time series; moving average terms are the lags are the lags of forecast errors; and an integrated version of a stationary series is a time series that is differenced to be made stationary (Nau, 2005).

For illustration, two basic ARIMA models are the ARIMA(1,0,0) and ARIMA(0,0,1). Equations 2.8 and 2.9 show the mathematical forms of these models:

$$\text{ARIMA(1,0,0): } Y_t = \mu + \varphi_1 Y_{t-1} + e_t \tag{2.8}$$

$$\text{ARIMA(0,0,1): } Y_t = \mu + e_t - \theta_1 e_{t-1} \tag{2.9}$$

where μ is a constant, φ is an autoregressive term, θ is a moving average term, and e is the error term. However, the ARIMA(1,0,0) and ARIMA(0,0,1) are equivalently AR(1) and MA(1) models, respectively, as autoregressive and moving average models are subsets of ARIMA.

General Seasonal ARIMA Model.

Some time series exhibit seasonal properties in addition to non-seasonal ARIMA characteristics. An extension of the non-seasonal ARIMA model accounts for seasonal aspects of time series. The notation for seasonal models is ARIMA(p, d, q)(P, D, Q)$_s$. Similarly, for an ARIMA(p, d, q)(P, D, Q)$_s$

- P is the number of seasonal autoregressive terms,
- D is the number of non-seasonal differences, and
- Q is the number of lagged forecast errors (Nau, 2005).

The seasonal aspects of time series appear in the ACF and the PACF. To determine seasonality, forecasters examine statistically significant lags in ACFs and PACFs (Makridakis, Wheelwright, & Hyndman, 1998).

Box-Jenkins Approach

Box and Jenkins describe a basic, three-phase approach to the development of an ARIMA model. The first phase of the Box-Jenkins approach is Identification. During Identification, forecasters prepare the data and select the model. The extent of data preparation depends on the characteristics of the time series that may or may not violate the stationary assumptions of the ARIMA model.

Autocorrelation functions (ACF), partial autocorrelation functions (PACF), and data characteristics influence model selection. Autocorrelation functions inform forecasters of the relationships between observations with distinct times of separation. A statistically significant autocorrelation at a specific lag indicates a potential time-dependency of a current observation on the observation at the time difference.

Similarly, partial autocorrelation functions measure the relationships between explanatory variables with various times of separation. The value of PACFs is that they

show transitive relationships. Particularly, if observations Y_t and Y_{t-1} have a significant correlation, Y_{t-1} and Y_{t-2} also have a significant correlation because the time difference is the same. Y_t and Y_{t-2} will have a correlation through the common relationships to Y_{t-1}. Partial autocorrelation measures the correlation of Y_t and Y_{t-2} with the removal of the intermediate Y_{t-1} observation (Makridakis, Wheelwright, & Hyndman, 1998).

The second phase of the Box-Jenkins approach is Estimation and Testing. Estimation involves determination of parameters and model rank criteria for potential models. Forecasters use model rank criteria to evaluate collections of parametric models with different numbers of variables. Akaike's Information Criterion (AIC) and Schwarz's Bayesian Criterion (SBC) are standard model rank criteria. Both criteria rank models using a tradeoff between model accuracy and model complexity. All else equal, the criteria favor parsimonious or terse models. We list the equations for AIC and SBC, respectively, in Equations 2.10 and 2.11:

$$AIC = -loglikelihood + 2k \tag{2.10}$$

$$SBC = -2loglikelihood + kln(n) \tag{2.11}$$

where k is the number of parameters (Akaike, 1974) (Schwarz, 1978). Testing, particularly diagnostics, determines if the chosen model meets the third assumption for an ARIMA model: the forecast errors are uncorrelated "white noise".

The final phase of the Box-Jenkins approach to ARIMA model development is Application. Simply, the intrinsic value of the ARIMA model lies in the performance of the model in-practice. Figure 2.3 summarizes the phases and elements of the Box-Jenkins approach to ARIMA model development.

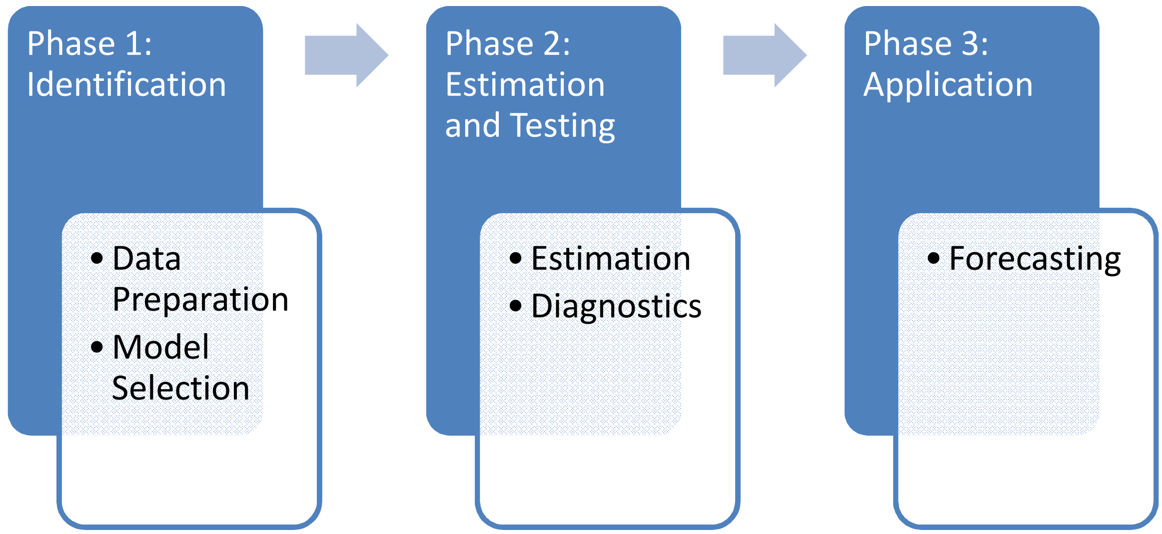


Figure 2.3: Box-Jenkins Approach to ARIMA Model Development (Makridakis, Wheelwright, & Hyndman, 1998)

Analysis of Contractor Cost Data

The *Guide to Analysis of Contractor Cost Data* provides guidance to acquisition analysts on the analysis of DoD contractor cost and schedule data (Headquarters Air Force Materiel Command, Financial Management, 1994). The intent of the guide is to aid acquisition programs in the reduction of cost growth and the improvement of visibility. The guide discusses numerous analytical techniques that focus on cost, schedule, and technical performance. Acquisition analysts study many of these measures (e.g. Cost Performance Index (CPI) and Schedule Performance Index (SPI)), in-practice.

Additionally, the guide offers guidance on the use of problem analysis techniques. Problem analysis techniques include measures of cost and schedule efficiency, variance verification, management reserve analysis, manpower loading trend analysis, performance trends, forecasting by Estimate at Complete (EAC) function, and Over

Target Baseline (OTB) analysis. One particularly useful metric is the percent complete versus percent spend chart, which shows the cost and schedule performance expectation. Figure 2.4 illustrates what constitutes a "normal" percent complete-percent spent chart. Deviations from the normal percent complete-percent spent line may indicate cost problems for an acquisition program. Similarly, Figure 2.5 shows a normal percent complete vs. percent scheduled chart. Deviations from the normal percent complete-percent scheduled chart may indicate problems for an acquisition program.

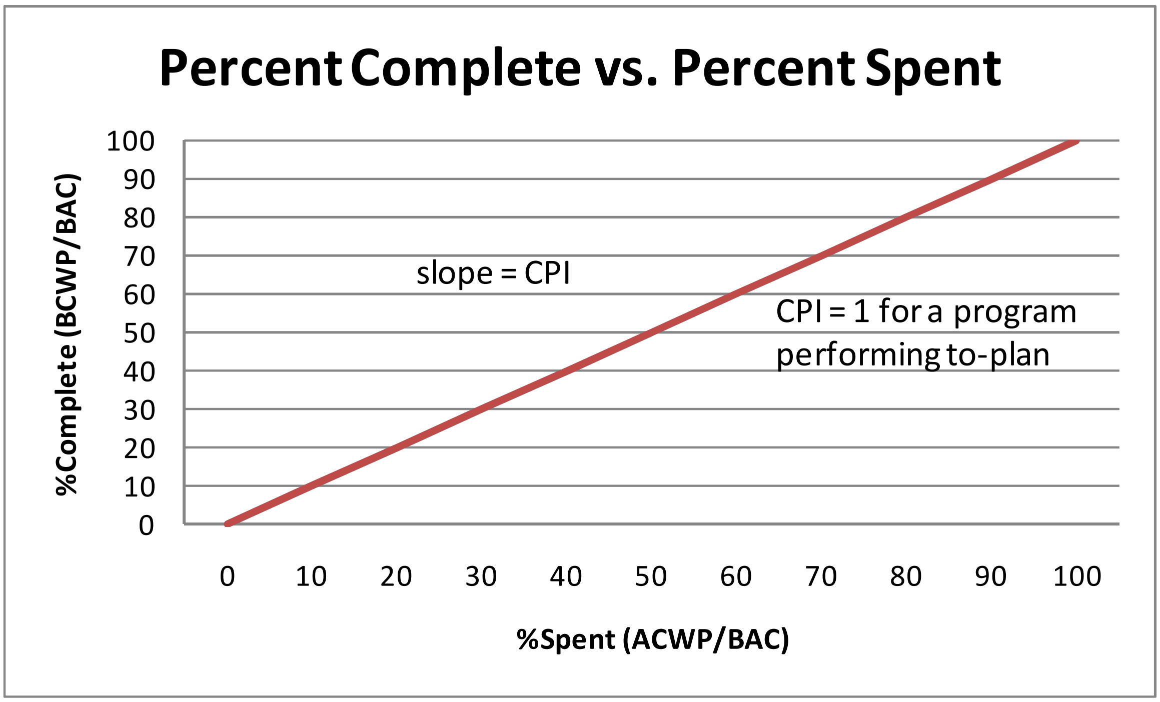


Figure 2.4: "Normal" Percent Complete vs. Percent Spent Chart

(Headquarters Air Force Materiel Command, Financial Management, 1994)

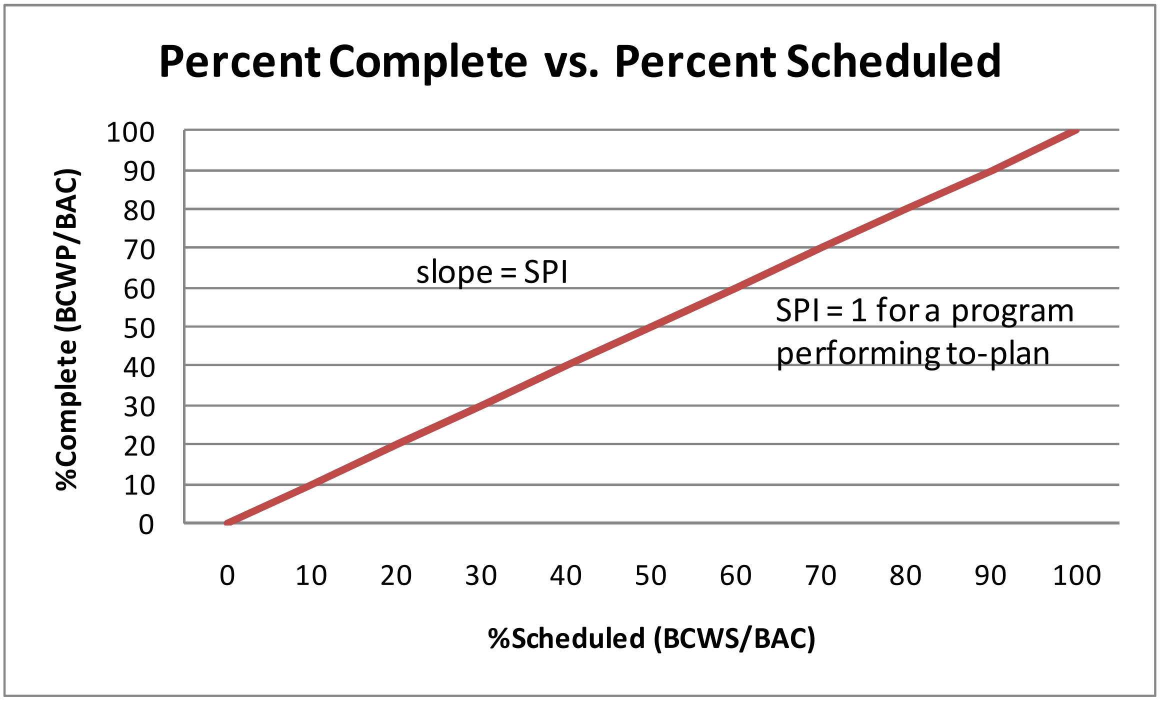


Figure 2.5: "Normal" Percent Complete vs. Percent Scheduled Chart

(Headquarters Air Force Materiel Command, Financial Management, 1994)

This chapter outlined change detection techniques, specifically the general class of ARIMA forecasting models. We discussed the assumptions and tests for ARIMA that ensure the accurate characterization of the data. Finally, we overviewed the percent complete-percent spent and percent complete-percent scheduled charts to illustrate what normal cost and schedule performance for an acquisition contract looks like. In the next two chapters, we apply ARIMA techniques to model earned value data. We use the model to develop an algorithm that detects changes from the normal value of 1 for the CPI and the SPI.

III. Methodology

This analysis studies online change detection of earned value (EV) data to identify and isolate potential problems in acquisition contracts. In this chapter, we discuss our approach to this change detection analysis. We begin with a description of the data source, our contract selection criteria, and the limitations of the data source. Next, we discuss the EV measures we select from the data source, our categorization process, and the normalization procedure for these measures. Finally, we 1) explain why and how we forecast EV data with ARIMA models, 2) describe our approach for detecting changes in the EV time series, and 3) compare change times to deviations in the percent complete vs. percent spent chart.

Data Source

The Defense Cost and Resource Center (DCARC) hosts a major collection of detailed EV data for Department of Defense (DoD) acquisition contracts. These data include monthly Contract Performance Reports (CPR), contract history files, and other EV and programmatic data submissions directly from program offices. For this analysis, we use EV history files available in DCARC.

Contract Selection Criteria.

We use contract history files because they contain panel data for fundamental earned value metrics. Specifically, contract history files include data for Actual Cost of Work Performed (ACWP), Budgeted Cost of Work Performed (BCWP), Budgeted Cost of Work Scheduled (BCWS), analytical derivatives of ACWP, BCWP, and BCWS (e.g.

Cost Variance (CV) and Schedule Variance (SV)), Estimate at Complete (EAC), Budget at Complete (BAC), Management Reserve (MR), categorical information, and report dates for each Work Breakdown Structure (WBS) element at all levels. Additionally, because DoD and the American National Standards Institute (ANSI) maintain specific requirements and instructions for these measures, we assume the data provide a framework for reliable measurement (OUSD(AT&L)ARA/AM(SO), 2005) (NDIA/PMSC, 2009).

We limit our analysis database to history files for Research, Development, Test, and Evaluation (RDT&E) contracts in DCARC. We select RDT&E contracts because typically they are large budget contracts with high cost and schedule uncertainty and risk. Alternatively, production contracts normally have less uncertainty and risk that may artificially eliminate the changes we wish to detect.

In an internal query of DCARC, we identify 813 files which meet our database specifications. Of the 813 contracts we identify in our information query, we locate only 787 files in the database. The different file types of the search results (e.g. .pdf and .trn) reduce the number of files we can access from 787 to 165 because we cannot extract all data automatically (i.e. without a major manual data entry effort). Finally, of the 165 files we can access, we find 32 unique contract history files for RDT&E contracts. We eliminate one history file due to data inconsistencies (Table 3.1). In Table 3.2 and Table 3.3 we list the number of contracts in the research database by Military Handbook Type and military service, respectively.

We do not impose a contract start date or end date constraint on the research database due to the small number of history files we gather from DCARC; however, the start date for all but one contract is after 1 January 2000 (Table 3.4).

Table 3.1: Database Size Reductions

Database Size Reductions	Number of Files
Search Results	813
Files Available	787
Accessible Files	165
Unique History Files	32
History Files in Research Database	31

Table 3.2: Number of Contracts by Military Handbook Type

Military Handbook Type	Number of Contracts
Aircraft	8
Electronic/Automated Software	13
Missile	3
Ship	1
Space	3
Surface Vehicles	2
System of Systems	1
Total	31

Table 3.3: Number of Contracts by Military Service

Military Service	Number of Contracts
Air Force	11
Army	7
Navy	12
Department of Defense	1
Total	31

Table 3.4: Number of Contracts by Contract Start Date

Contract Start Date	Number of Contracts
1 Jan 1995-31 Dec 1999	1
1 Jan 2000-31 Dec 2004	11
1 Jan 2005-31 Dec 2009	19
Total	31

Limitations of Data Source.

In reality, we use a data source with an unintentional filter for this analysis. The data source is the collection of acquisition contract history files; the filter is DCARC. The result of the collection-filter process is a smaller pool of contract history files.

Acquisition contract history files offer some benefits, but pose many obstacles to analysis. The principle benefit of contract history files is that they provide time series data at multiple levels of the contract WBS. The obstacles are three-fold. First, a contract history file is effectively a concatenation of sequential monthly CPRs. Often monthly CPRs contain inaccuracies which program offices work with the contractor to correct. CPR re-submissions to DCARC are evidence of this issue. However, in some instances systematic errors persist in the contract history files we collect. We attempt to resolve these data issues with the appropriate monthly CPRs or the applicable CPR resubmissions.

Second, a contract history file does not always contain the full time series. One reason for partial time series is many program offices update their contract history files on an annual basis. Thus, a researcher who collects history files between updates may not acquire the additions to the time series since the last release. In Appendix E, we show the percentage of the total contract that each of the contracts in the research database covers. We calculate percent coverage by comparing the contract start date and contract end date to the available months of data in the contract history files.

Third, the flexibility in electronic submission formats permitted by the CPR-governing Data Item Description (DID) creates data accessibility issues for cross-program analysis that individual program offices may not face

(OUSD(AT&L)ARA/AM(SO), 2005). Specifically, our data processing and management resources cannot process all file types that contractors submit. Individual program offices likely do not have this issue because they have a direct relationship with the contractor and can specify an electronic format both can handle easily.

The main limitation DCARC imposes on our research database is the file size that program offices can upload to the database. Although we do not encounter this problem directly, indirectly file sizes that are too large to submit are unavailable in DCARC and thus impact the number of contract history files we collect. As a result, DCARC inadvertently filters available contract history files.

Another limitation of DCARC is the number of months of data available for each contract. Generally, the length of the time series in a contract history is shorter than the time from contract start date to present. Thus, some of the contract history files we use have fewer months than the contract's actual number of months to-date.

Earned Value Data

We construct our research database with entries for ACWP, BCWP, and BCWS with respect to report date for each contract history file. We sort these using WBS level as the criterion.

Categorization.

For the WBS level criteria, we sort the data by level 1 and sum the values within the level. These sums are cumulative values for ACWP, BCWP, and BCWS. We limit the sort criteria to WBS level 1, but conceivably can use level 2 and 3 also. Data for WBS levels greater than 3 are problematic because fewer contracts report at each lower

level and thus reduces the sample size increasingly. Different sample sizes create data comparison issues between acquisition contracts.

We compute monthly ACWP, BCWP, and BCWS values and monthly and cumulative analytic earned value measures for the level 1 data. The analytic EV measures we calculate are:

- Cost Variance (CV$)
- Normalized Cost Variance (NCV)
- Percent Cost Variance (%CV)
- Schedule Variance (SV$)
- Schedule Variance (SVMonths)
- Normalized Schedule Variance (NSV)
- Percent Schedule Variance (%SV)
- Cost Performance Index (CPI)
- Schedule Performance Index (SPI)
- To-Complete Index (TCPI).

The equations we use to calculate the analytic EV measures are shown in Appendix A.

Data Normalization.

Differences in the size (e.g. Budget at Complete (BAC)), contract length, and inflation can complicate comparisons among contracts. We address how we deal with these issues of contract comparability.

First, the importance of a change in ACWP, BCWP, or BCWS is relative to the size of the contract. Although a change may be large in amount, the relative change may be small compared to the size of the overall contract. However, calculations for CPI and SPI control for contract size because changes in ACWP, BCWP, and BCWS are relative to one another.

Next, the length of a contract may influence how abruptly a change appears over an entire contract. Traditionally, EV analysts use a percent complete calculation to

manage this concern. In this analysis, we focus on monthly changes, not changes throughout entire contracts. Therefore, contract length does not affect our analysis.

Finally, the effect of inflation creates disparities in the value of money across time. We use 2010 as a base year (BY10$) to standardize costs in time. We gather the conversion rates from the 2010 release of Deputy Assistant Secretary of the Air Force for Cost and Economics (SAF/FMC) inflation tables (SAF/FMC, 2010).

Forecasting Earned Value Data with ARIMA Models

ARIMA forecasting offers a logical approach to online change detection in earned value data. We theorize patterns in cumulative ACWP, cumulative BCWP, and cumulative BCWS time series are distinguishable from data noise. We can model these patterns to determine how we can best show real-time changes in the CPI and the SPI. Although we lack a large amount of data for any single program, our database has enough observations to confirm trends for several programs. Lastly, we expect historic cost and schedule performances to continue in the future.

We analyze our time series in JMP® version 9. The time series capability in JMP® includes ARIMA models which we use to forecast EV data. The parameter test statistics and rank criteria we obtain from JMP® help us appraise each acquisition contract model in our research database. We record consistent time series characteristics to consider during model selection.

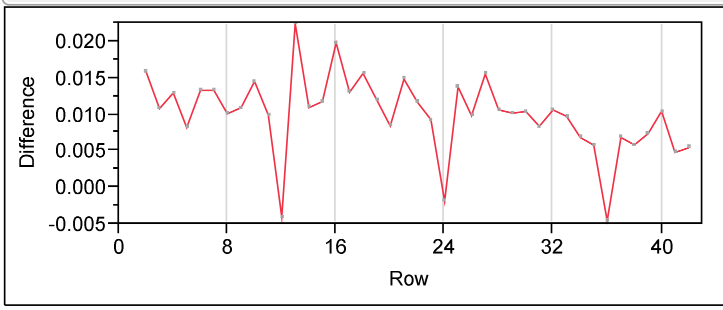
Largely, we conduct our analysis using the Box-Jenkins approach. We begin with plots of the time series for each acquisition contract. We plot each time series to examine if the means and variances are stationary for the ACWP, BCWP, and BCWS time series.

We find the means of the time series are non-stationary and require differencing. The variances of the time series are stationary and thus do not require transformation.

JMP® plots of the differenced time series (Figure 3.1), autocorrelation functions (ACF) (Figure 3.2, left), and partial autocorrelation functions (PACF) (Figure 3.2, right) allow visual verification that the differenced time series are stationary. The data used to plot Figure 3.1 and Figure 3.2 is an example of a differenced time series from the research database. Figure 3.1 indicates the mean and the variance are stationary because the data are distributed about a constant mean without a growing or decaying variance. Despite the potential pattern shown by the recurrence of dips at regular intervals, no hypothesis test indicates a significant change in the mean, likely because the number of observations reduces the power of the test. Additionally, the ACF and PACF plots in Figure 3.2 show the mean is stationary because the values reduce to zero quickly.

The Augmented Dickey-Fuller test (ADF) illustrates mathematically the time series are stationary. Specifically, to reject the null hypothesis at some level of confidence, the ADF must be a negative value, with greater negativity reflecting a higher level of confidence. The ADF values in Figure 3.1 for zero mean, single mean, and trend confirm the time series are stationary because the values are all negative.

Difference: (1-B)^1

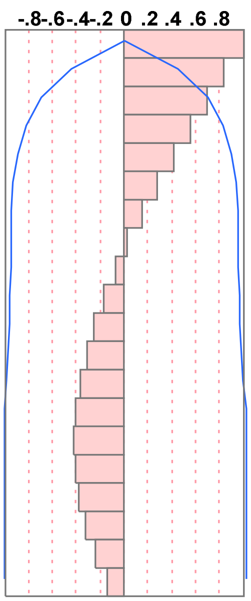


Mean	0.0100318
Std	0.0052385
N	41
Zero Mean ADF	-2.332796
Single Mean ADF	-5.911811
Trend ADF	-6.882461

Figure 3.1: First Non-Seasonal Difference

Time Series Basic Diagnostics

Lag	AutoCorr	Ljung-Box Q	p-Value	Lag	Partial
0	1.0000	.	.	0	1.0000
1	0.8460	16.5747	<.0001*	1	0.8460
2	0.7010	28.5868	<.0001*	2	-0.0518
3	0.5573	36.6246	<.0001*	3	-0.0800
4	0.4145	41.3488	<.0001*	4	-0.0888
5	0.2780	43.6161	<.0001*	5	-0.0781
6	0.1497	44.3201	<.0001*	6	-0.0767
7	0.0292	44.3490	<.0001*	7	-0.0825
8	-0.0730	44.5444	<.0001*	8	-0.0509
9	-0.1652	45.6364	<.0001*	9	-0.0767
10	-0.2483	48.3486	<.0001*	10	-0.0847
11	-0.3138	53.1624	<.0001*	11	-0.0594
12	-0.3671	60.5759	<.0001*	12	-0.0737
13	-0.4058	70.9243	<.0001*	13	-0.0655
14	-0.4178	83.7239	<.0001*	14	-0.0184
15	-0.4093	98.4668	<.0001*	15	-0.0200
16	-0.3768	114.081	<.0001*	16	0.0072
17	-0.3195	129.057	<.0001*	17	0.0321
18	-0.2409	141.821	<.0001*	18	0.0507
19	-0.1382	150.223	<.0001*	19	0.0895



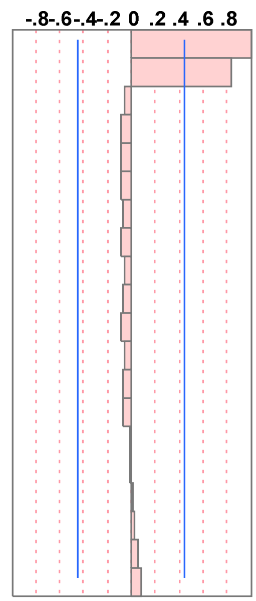


Figure 3.2: Plots of ACF and PACF

The ACF and PACF plots also reveal potential autoregressive (AR) models, moving average (MA) models, or seasonality. Bars that exceed the boundary lines in the ACF or PACF indicate statistically significant lags. In Figures 3.1 and 3.2, we find alternative representations of a statistically significant lag of 1. This lag of one period implies an observation one period earlier in the time series influences the current

observation. As a result, AR(1) and MA(1) models may be appropriate. However, we observe no statistically significant lags at seasonal intervals.

We observe upward trends and lag 1 characteristics in the ACWP, BCWP, and BCWS time series for all contracts, but we do not observe seasonal patterns. Therefore, we confine our model selection to non-seasonal ARIMA models that account for these characteristics. We use the ARIMA model group function in JMP® to test models that meet the inclusive range of specifications for p, d, and q in Table 3.5. We identify eight potential models for the combination of these p, d, and q ranges. Table 3.6 lists these eight models.

Table 3.5: Bounds ARIMA Model Characteristics

ARIMA	Minimum	Maximum
p	0	1
d	0	1
q	0	1

Table 3.6: Potential ARIMA Models

Number	ARIMA Model
1	ARIMA(0,0,0)
2	ARIMA(1,1,1)
3	AR(1)
4	ARI(1,1)
5	ARMA(1,1)
6	I(1)
7	IMA(1,1)
8	MA(1)

The ARIMA model group function ranks models by the Akaike Information Criterion (AIC) and Schwarz Bayesian Criterion (SBC). The smaller the AIC and SBC values, the better rank the model earns (Akaike, 1974) (Schwarz, 1978). The rank

structure provided the ARIMA model group function was consistent between AIC and SBC measures. We find a reliable division in the potential models group with the R^2 values for the models. This division is important because R^2 measures the extent that the terms in the model explain the variation of the forecast. The few terms in all potential models alleviates concerns of overfitting the time series and, thus, the benefit of using adjusted R-square as a measure of model performance instead of R-square.

The division in potential models separates ARIMA models ARI(1,1), IMA(1,1), I(1), and ARIMA(1,1,1) from AR(1), MA(1), ARIMA(0,0,0), and ARMA(1,1). Table 3.7 lists the number of contracts in which each model occurred in the top four ranks according to the AIC and SBC measures. Because the first four models listed appear in the top four model ranks for nearly every program, we choose to examine these models further. We note two contracts have time series models that are not present in any other contract's top four ranks. The models that appear in these contracts' top four ranks are ARIMA(0,0,0) and AR(1). We believe they occur in the top four ranks because the two contracts have small numbers of observations.

Table 3.7: Number of Top Four Occurrences by AIC and SBC

ARIMA Model	Contracts		
	ACWP	BCWP	BCWS
ARI(1,1)	31	31	31
IMA(1,1)	30	30	30
I(1)	30	30	30
ARIMA(1,1,1)	30	30	30
ARIMA(0,0,0)	2	2	2
AR(1)	1	1	1
MA(1)	0	0	0
ARMA(1,1)	0	0	0

We validate the appropriateness of the high-occurrence model group [ARI(1,1), IMA(1,1), I(1), and ARIMA(1,1,1)] with tests of statistical significance for the terms in each model. Table 3.8 lists the number of contracts in which all parameters for a given model are statistically significant (α = 0.05). We find three out of four models in the high-occurrence group have one or more variables that are not statistically significant for approximately half of the contracts in the research database. Again, we find the same two contracts that have uncommon ARIMA models reduce the number of statistically significant models.

Table 3.8: Contracts with Statistically Significant Parameters (α = 0.05)

ARIMA Model	Contracts		
	ACWP	BCWP	BCWS
I(1)	28	27	27
IMA(1,1)	16	11	9
ARI(1)	13	11	11
ARIMA(1,1,1)	10	9	10

The I(1) model performs well against the model rank criteria and passes the tests of statistical significance for nearly all contracts. For this reason, we discard the other models and test the normality of residuals for the I(1) model only.

We use the Shapiro-Wilk and Ljung-Box methods to test the normality of the residual distributions. With the Shapiro-Wilk test, we standardize all residual values so we can compare the cumulative ACWP, cumulative BCWP, and cumulative BCWS time series to the normal distribution for all programs simultaneously. Table 3.9 reports the results of the Shapiro-Wilk normality test for the cumulative ACWP, cumulative BCWP, and cumulative BCWS time series. We reject the null hypothesis that the normal distribution approximates the distributions of the residuals for all time series (α = 0.05).

Although visually, we find the residuals are clustered closely around zero (see Appendix B for the distributions of the standardized residuals). We identify two residuals are outliers because they are further than three standard deviations from the mean. These outliers are approximately 6.5 and -4.0 standard deviations away from the mean.

We locate these outliers in our research database to examine why our time series model performs poorly on their prediction. Although the two outliers occur in different contracts, we find a common characteristic in the months that immediately precede the months of the outliers. Specifically, the months that precede both outliers have increasingly narrow forecast confidence intervals because sequential values for the cumulative ACWP, cumulative BCWP, and cumulative BCWS show precise monthly ACWP, BCWP, and BCWS performance rates. As a result, the forecast confidence interval narrows with each new month's data and thus even minor deviations from the monthly rates appear major.

Table 3.9: Shapiro-Wilk test of residuals ($\alpha = 0.05$)

Time Series	Fail to Reject	Reject
ACWP		X
BCWP		X
BCWS		X

For our second test of residual normality, we compare the Ljung-Box Q-value at lag 1 to critical values of the chi-square distribution. We use one lag period because this is the longest lag we consider in model selection. We evaluate the Q statistic with different degrees of freedom because the contracts span different numbers of months. Table 3.10 lists the results of Ljung-Box portmanteau test residuals ($\alpha = 0.05$). We fail to

reject the null hypothesis that the time series are normally distributed for 29 of 31 contracts. We reject the null hypothesis for the two previously-noted unusual time series.

Table 3.10: Ljung-Box Portmanteau Test of Residuals ($\alpha = 0.05$)

Result	Contracts		
	ACWP	BCWP	BCWS
Fail to Reject	29	29	29
Reject	2	2	2

Of the two tests for residual normality, the Shapiro-Wilk test is more mathematically robust because the Ljung-Box method sometimes fails to reject models that fit the normal distribution to the time series residuals poorly (Makridakis, Wheelwright, & Hyndman, 1998). As a result, we do not achieve the theoretical result of normally distributed residuals for the cumulative ACWP, cumulative BCWP, and cumulative BCWS time series.

However, theoretical data is often much different than actual data. Due to this difference, we attempt to characterize the clustered nature of the residuals more generally. Specifically, we determine if the true mean of the residuals falls within a certain confidence interval. If the residuals fall within a specific confidence interval, we can describe the statistical boundaries of the residuals for any distribution.

Chebychev's Theorem specifies the percentage of observations that fall within a confidence interval $\mu \pm k\sigma$ regardless of the distribution, where μ is the mean, σ is the standard deviation, k is the number of standard deviations such that $k > 1$.
(Newbold, Carlson, & Thorne, 2010). The Theorem states for any population the percent of observations that fall within the confidence interval is at least

$$100\left(1-\left(1/_{k^2}\right)\right)\%. \tag{3.1}$$

Although Chebychev's Theorem offers a practical method to guarantee a confidence interval for any population, the main limitation of the Theorem is the level of confidence for many populations is greater than result from Equation 3.1 (Newbold, Carlson, & Thorne, 2010).

We apply Chebychev's Theorem to the distributions of residuals for the ACWP, BCWP, and BCWS time series. By using the Theorem, we tradeoff more precise confidence levels for a theoretical minimum confidence level. We exclude the two statistical outliers from our calculations of the mean and the standard deviation for each time series (see Appendix C). Additionally, because we use plus or minus three standard deviations from the mean, according to Equation 3.1, the true value of the mean lies in the confidence interval in least 88.9% of all analyses. We list the confidence intervals for the distributions of residuals in Figure 3.11. The Lower Confidence Limit (LCL) and Upper Confidence Limit (UCL) are the boundaries of the interval. With exception of the two statistical outliers, we find no residuals outside these intervals for the population.

Table 3.11: Confidence Intervals for Standardized Residuals (CL = 88.9%)

Time Series	μ	σ	*k*	LCL	UCL
ACWP	-0.002351	0.958000	±3	-2.876351	2.871649
BCWP	-0.002421	0.957468	±3	-2.874825	2.869983
BCWS	-0.002343	0.957896	±3	-2.876031	2.871375

The characterization of the ACWP, BCWP, and BCWS time series with confidence intervals supports the modeling of the CPI and the SPI time series because ACWP, BCWP, and BCWS are inputs to the CPI and SPI. We use an ARIMA(0,0,0) or

"white noise" model for the CPI and the SPI. An ARIMA(0,0,0) is an appropriate model because we expect contracts with normal cost and schedule performances to have CPIs and SPIs equal to 1 and the normal distribution to characterize the error terms. We do not need to stationarize the mean or the variance because both are stable in the time series. Additionally, there is no requirement to test the statistical significance of the parameters because an ARIMA(0,0,0) only includes the error term.

We evaluate the normality of the residuals with the Shapiro-Wilk test to ensure the normal distribution models the error terms of the CPI and the SPI. Even though the means of the distributions of the residuals are approximately centered on zero and residual observations decrease away from the mean, the residuals fail to meet the assumption of normality (see Appendix D). However, because the distribution of the standardized residuals is robust against deviations from normality provided the distribution is relatively symmetric, we assume normality for both time series.

Change Detection

We use statistical differences to monitor real-time changes in the monthly Cost Performance Index (CPI) and Schedule Performance Index (SPI) observations. We theorize changes in the CPI and the SPI may indicate contract problems because these measures are the slopes of the percent complete vs. percent spent and percent complete vs. percent scheduled plots, respectively. We define a difference as a CPI or a SPI value statistically different from 1.

We use the Chebychev confidence interval in Equation 3.2 to specify the uncertainty boundaries for our forecast.

$$\bar{x} - ks < 1 < \bar{x} + ks \tag{3.2}$$

where $\bar{x}$ is the sample mean, s is the sample standard deviation, and k is the number of sample standard deviations (Newbold, Carlson, & Thorne, 2010). We use the sample mean and sample standard deviation because we "acquire" the observations we evaluate serially. We test the sensitivity of the algorithm for a series of standard deviations to tradeoff false detections (Type I errors) with missed detections (Type II errors); specifically, we test standard deviations from 0.5 to 3.0.

For example, when a standard deviation of 0.5 is used for the confidence interval the algorithm favors false detections in lieu of missed detections. The propensity towards false detections is because the probability density function (PDF) for one standard deviation of a normal distribution captures 38.2% of the distribution. Therefore, given observations data up to $\hat{y}_t$, the probability a forecast $\hat{y}_{t+1}$ is determined to be statistically different from the sample mean is 61.8% (α = 0.618). Plainly, about three-fifths of observations will "detected" as statistically significant changes.

For an accurate estimate of the standard deviation, we do not begin change detection until the fourth observation. That is, the first observation for which we attempt to detect a change in each time series is the fourth month's observation. Theoretically, we can detect change with one and two observations used to calculate the standard deviation. Practically, however, we choose three prior observations to estimate the time series standard deviation with the expectation of a narrower confidence interval.

We compare months that indicate changes in the CPI or the SPI with months of major changes in contractor EACs. We theorize months that indicate change in our detection algorithm will lead or correspond to major changes in the contractor EAC. A change in the contractor EAC is a significant event because the company under contract acknowledges formally it likely cannot complete the work required at or within the dollar value of the current EAC.

We define major changes in the EAC as:

1. $\%\Delta EAC \geq 10\%$
2. $10\% > \%\Delta EAC \geq 5\%$
3. $-10\% < \%\Delta EAC \leq -5\%$
4. $\%\Delta EAC \leq -10\%$

We choose these categories to characterize major EAC changes because changes within 5% occur frequently and therefore likely represent normal data noise. Changes of at least 5% appear much less frequently and thus we theorize are indicative of major performance changes.

In this chapter we overviewed the data we used in this analysis and the limitations of the data. We explained how we modeled and tested the ACWP, BCWP, and BCWS time series. Finally, we discussed how we detect changes in the CPI and the SPI. In the next chapter, we review the results of the change detection analysis.

IV. Results and Discussion

In this chapter, we review the results of our change detection algorithm. We discuss the changes we detected and the characteristics of these changes. We explore in detail the time relationships of change detections and major changes in the Estimate at Complete (EAC).

Overall, we found 99 months had major percentage changes in the EAC out of 1094 potential months. Logically, the number of changes detected in the CPI and SPI increased with greater algorithm sensitivity. For perspective, the most sensitive algorithm we tested (0.5 standard deviations) identified 550 and 549 changes in the CPI and SPI, respectively. This algorithm sensitivity detected changes in approximately half of the 1094 observations in the research database and about five times the number of major EAC changes that occurred during the same month as the detections. The least sensitive algorithm (3.0 standard deviations) detected statistical changes in the CPI and SPI for 89 and 75 observations, respectively. Therefore, the least sensitive algorithm we tested detected changes in less than 10% of observations and less than 80% of the number of major EAC changes that occurred during the same month as the detections.

We observed a noticeable increase (approximately 30%) in detections of major EAC changes between the 0.5 and 1.0 standard deviation sensitivities (Figure 4.1). We contend the increase in detections was sensible because the probability density function (PDF) for a normal distribution at 0.5 standard deviations rejects the null hypothesis for 30% more observations than for 1.0 standard deviation. In other words, the algorithm

with 0.5 standard deviations will detect changes in about 2/3 of observations where the algorithm with 1.0 standard deviation will detect changes in about 1/3. We did not test algorithm sensitivities greater than 0.5 standard deviations and thus cannot inform the reader of the performance of algorithms more sensitive than the 0.5 standard deviation sensitivity. However, we expect algorithms more sensitive than 0.5 standard deviations will detect yet greater percentages of major EAC changes.

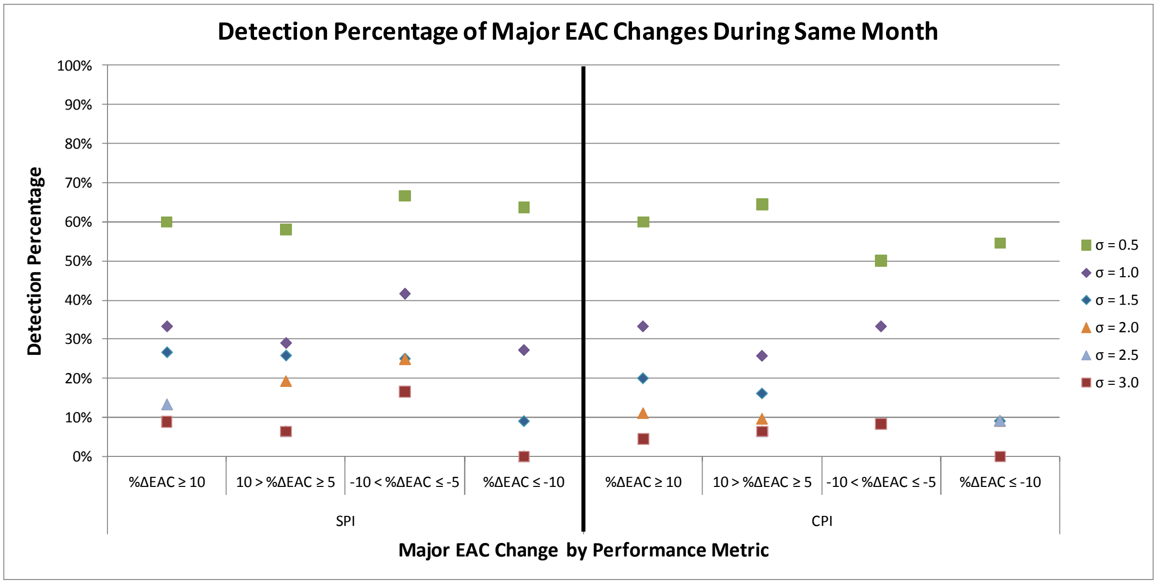


Figure 4.1: Detections and Major EAC Changes During Same Month

By algorithm sensitivity, Figure 4.2 illustrates the percentage of false detections of major EAC changes that during the same month as the detections. We calculated these percentages as the number of detections for non-major EAC changes relative to the total number of detections. In Figure 4.2, negative percentages reflect algorithm sensitivities that had smaller numbers of detections than major EAC changes in the same observation period. We find the greater the sensitivity of the algorithm, the greater the percentage of missed detections.

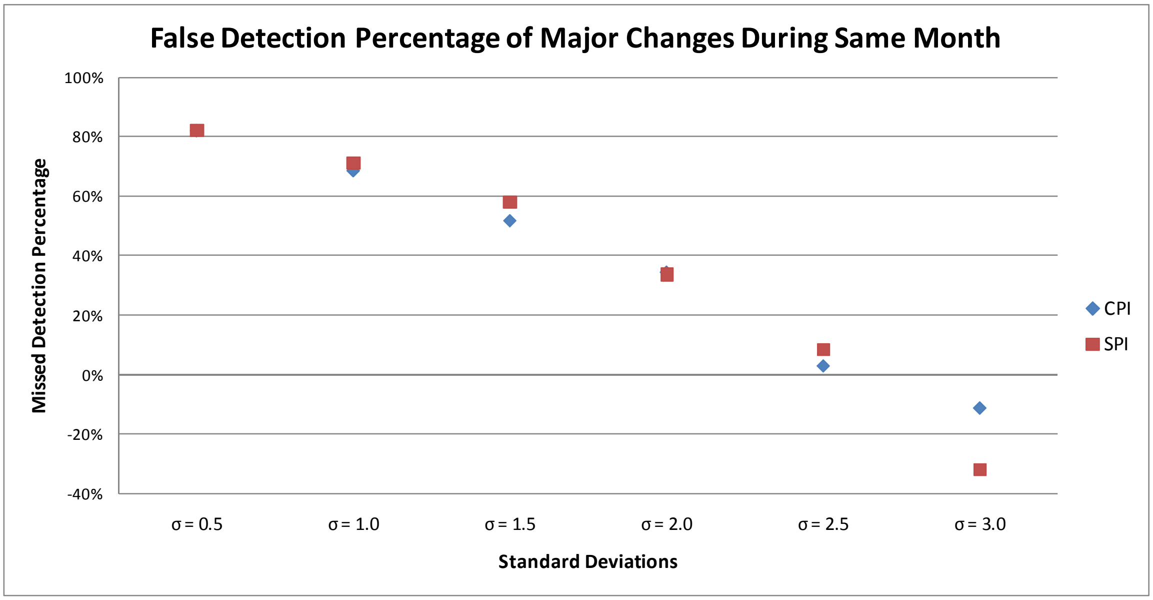


Figure 4.2: False Detections and Major EAC Changes During Same Month

In combination, Figures 4.1 and 4.2 show the tradeoff between algorithm sensitivity and detections for major EAC changes that occurred during the same month as the detections. Specifically, the more sensitive the algorithm, the higher the percentages of correct (or true) detections and the higher the percentage of false detections (Type I error). As algorithm sensitivity decreased, the percentage of correct detections also decreased while the percentage of missed detections increased (Type II error).

Figures 4.3 and 4.4 illustrate the tradeoff between algorithm sensitivity and detections for major EAC changes that occurred during months following the detections. We calculated the mean percentages for the values in Figure 4.3 to compress the time component of the detections which spanned 12 to 1 month(s) before major changes in the EAC. Again, we note the more sensitive the algorithm, the higher the percentages of correct and false detections. The less sensitive algorithms had lower correct detection percentages and higher missed detection percentages.

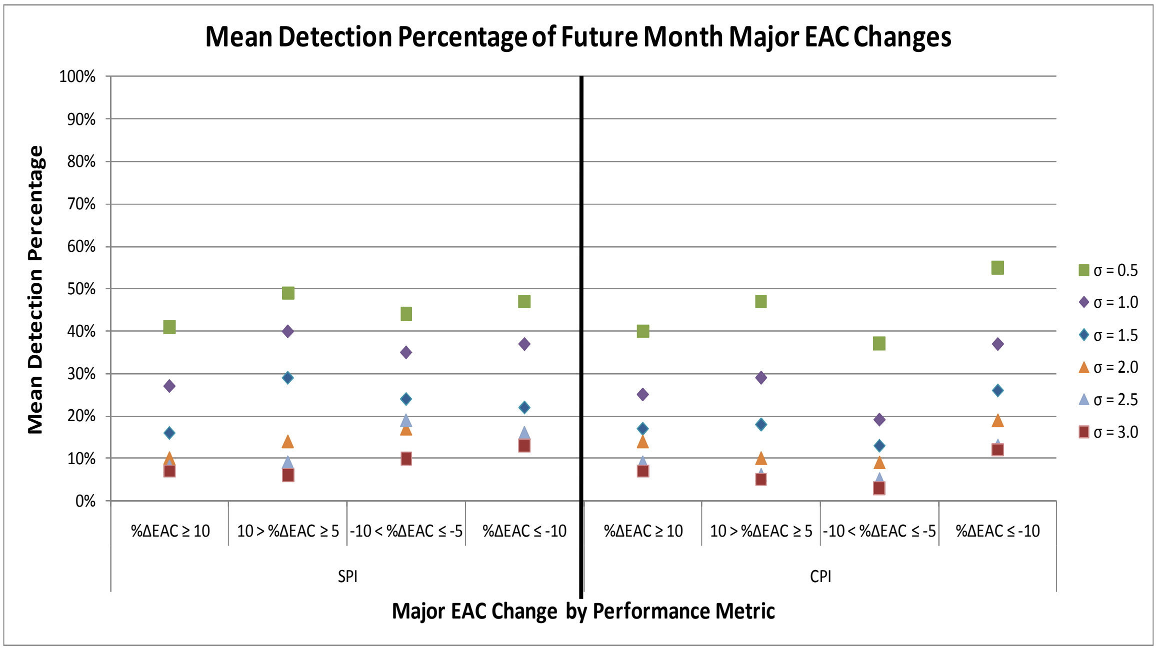


Figure 4.3: Mean Detection Percentage of Future Month Major EAC Changes

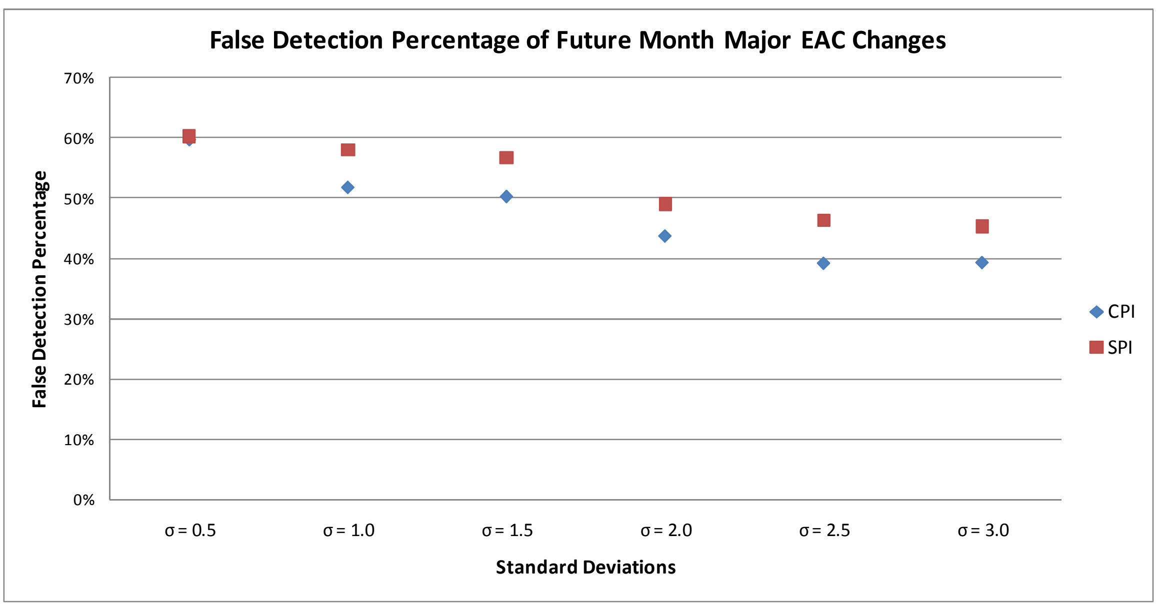


Figure 4.4: False Detection Percentage of Future Month Major EAC Changes

Frequency

Figures 4.5 and 4.6 show the percentages of changes detected in the CPI and the SPI for different standard deviations. For both CPI and SPI, observations exceeded the Lower Confidence Limit (LCL) more frequently than the Upper Confidence Limit (UCL): 83% and 84%, respectively. The higher percentage of LCL detections does not imply the algorithm is more sensitive to worsening cost and schedule performances. Rather, the algorithm does detect worsening cost and schedule performances, and in the database a higher ratio of worsening than improving performance was detected.

Current Month Detections

Figures 4.7 and 4.8 illustrate the percentages of major current month EAC changes detected by the algorithm. Intuitively, for increasing sensitivity—fewer standard deviations—the algorithm detects higher percentages of changes in the EAC for the current month. Likewise, for decreasing detection sensitivity, the algorithm does not detect higher percentages of changes in EAC.

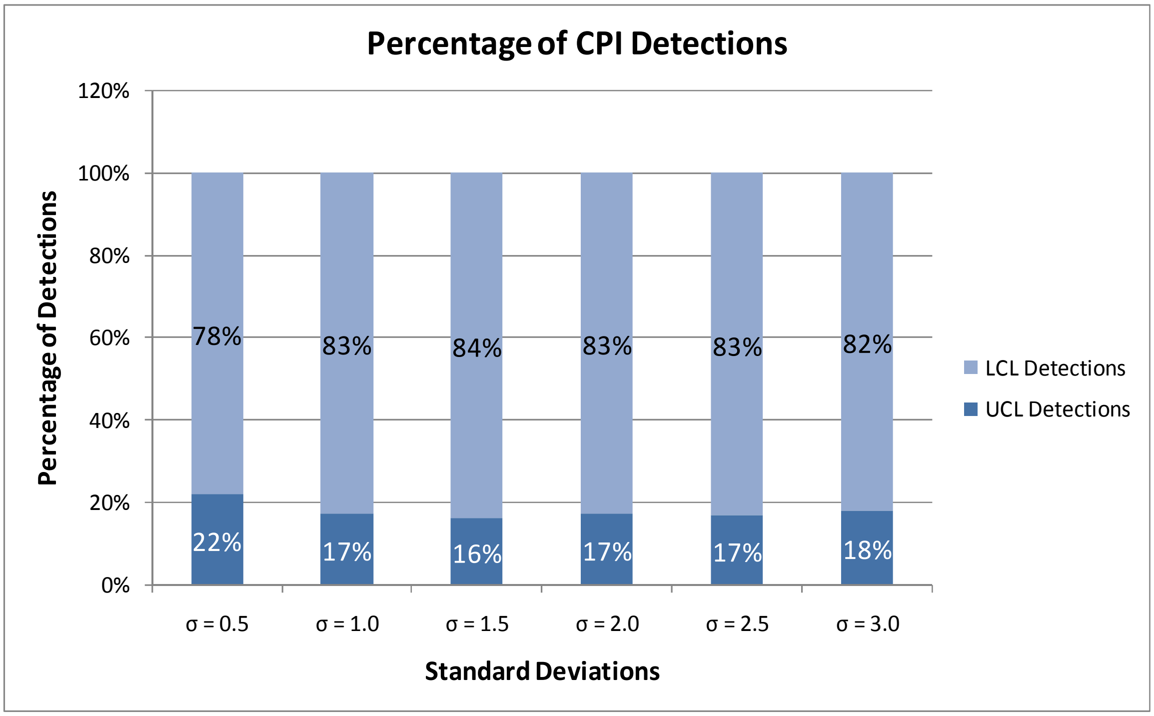


Figure 4.5: Percentage of CPI Detections

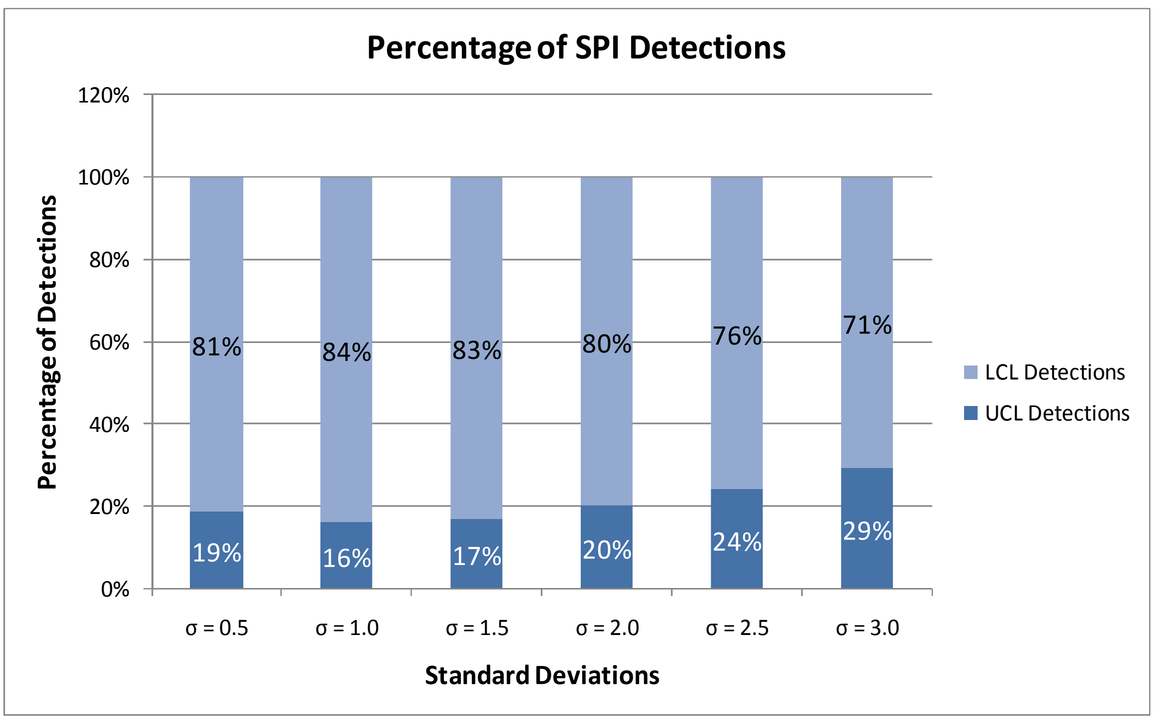


Figure 4.6: Percentage of SPI Detections

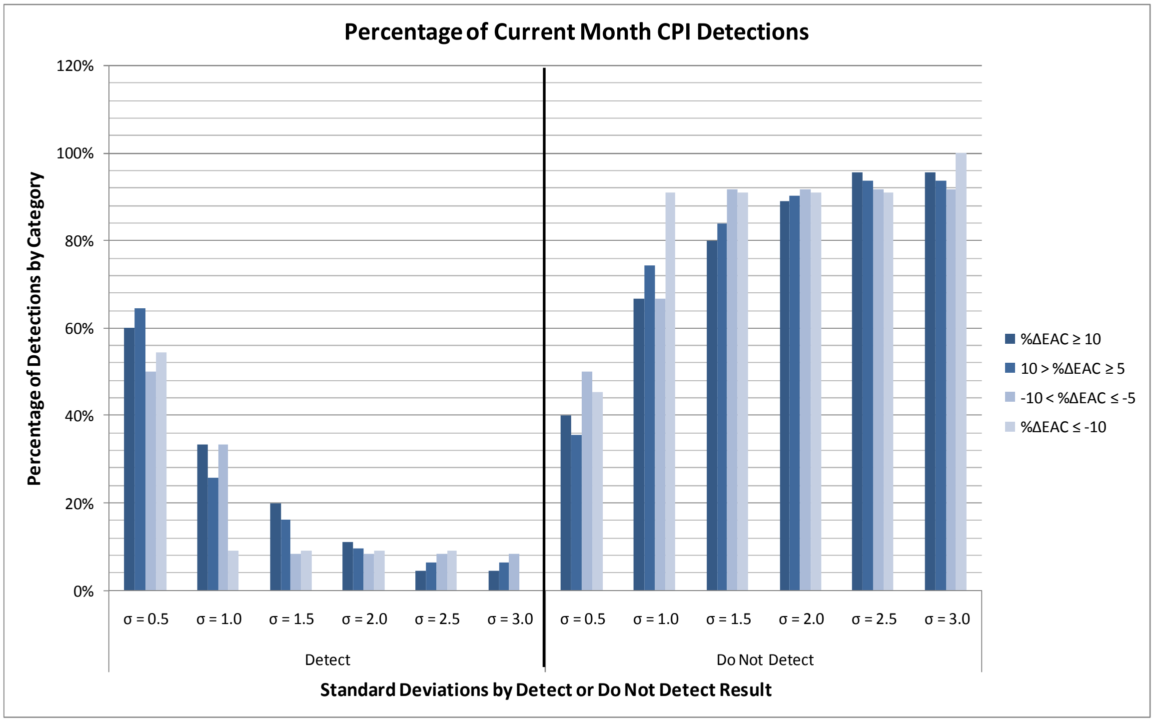


Figure 4.7: Percent age of Current Month CPI Detections

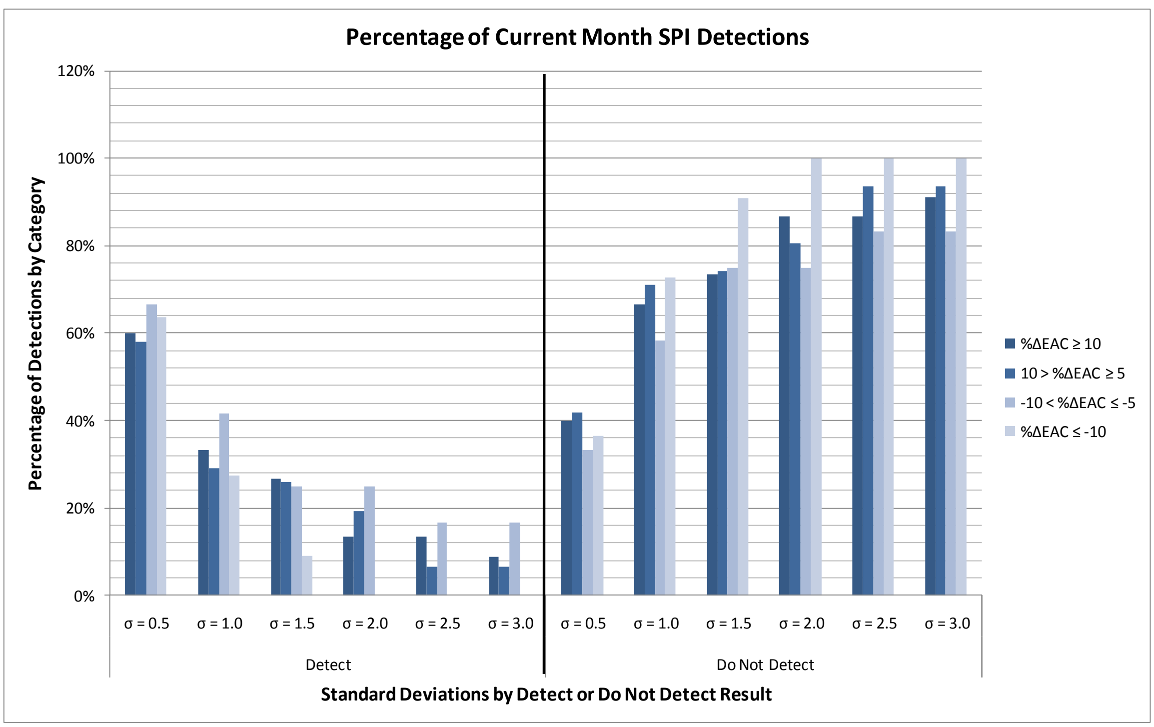


Figure 4.8: Percentage of Current Month SPI Detections

Early Detections

The algorithm identified informational early detection relationships between CPI or SPI detections and all groups of major EAC changes. Changes in the CPI and SPI corresponded to major changes in the EAC as early as twelve months before the EAC change. The percentage of detections grew as the time difference between the CPI or SPI detection decreased from the EAC change. Similarly, the number of non-detections decreased as time between detection and EAC change decreased (Figures 4.9 and 4.10). Although upward and downward trends are evident in Figures 4.9 and 4.10, clearly there are deviations from these overall trends which cause the trends to be jagged or unsmooth. We attribute these deviations to the small sample size.

Detection Relationships

The algorithm identified 185 occurrences of simultaneous CPI and SPI changes during the same month as a major change in the EAC. Of the 185 occurrences, 13 corresponded to major changes in the EAC (93% false detection rate). All major changes in the EAC were detected (0% missed detection rate). Table 4.1 lists the numbers and percentages of detections by group of major EAC change. We see 54% of the contracts experience at least 10% increases in EACs when there were simultaneous detections.

Table 4.1: Simultaneous CPI and SPI Detections During Same Month

	Same Month Detections	
% Change in EAC	# Detections	% Detections
EAC ≥ 10	7	54%
10 > EAC ≥ 5	3	23%
-10 < EAC ≤ -5	2	15%
EAC ≤ -10	1	8%
Total	13	100%

We examined the relationship between sequential detections for the CPI and SPI and a subsequent major change in the EAC. Specifically, we analyzed whether a detection in the CPI or the SPI was followed by a detection in the opposite index (CPI or SPI) during the next twelve months. If a sequential detection was identified, we looked for a major change in the EAC during the twelve months after the second detection; we found no such occurrences.

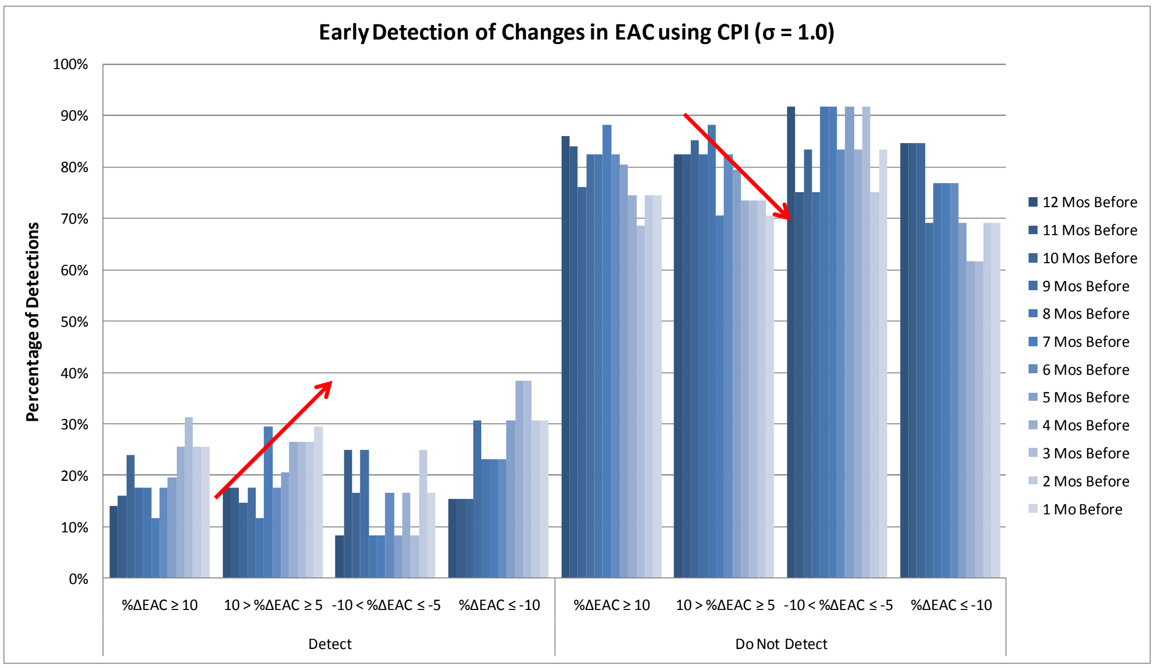


Figure 4.9: Early Detection of Changes in EAC Using CPI

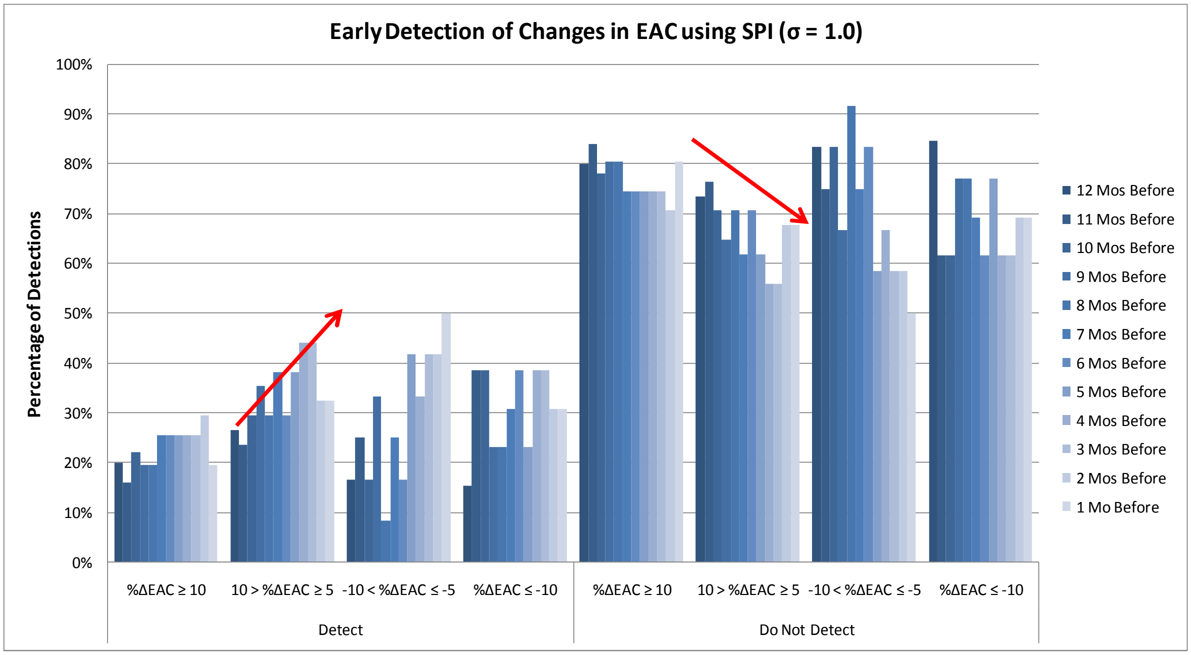


Figure 4.10: Early Detection of Changes in EAC Using SPI

In this chapter we reviewed the results of the change detection analysis. We found changes in the CPI and the SPI correspond to changes in the EAC during the same month and future months. The percentage of detections that correspond to major EAC changes increases as the length of time between the two decreases. We observed the detection of changes in the CPI and the SPI simultaneously corresponded to major increases in the EAC in 77% of occurrences. We did not see any relationship between delayed detections of the CPI or the SPI. In the final chapter, we summarize our results, discuss policy implications, and offer suggestions for further research.

V. Conclusions

In this chapter, we remind ourselves of the questions we sought to answer:

1. Can we detect changes in acquisition contracts with a detection algorithm?
2. If we can detect changes, how long does a problem exist before we identify it?

Review of Results

Our analysis of earned value data reveals we can detect changes in acquisition contract performance. We developed an algorithm based on an updating confidence interval to detect these changes. We found the change detection algorithm identifies worsening more often than improving cost and schedule performances. This result reflects the observations from prior contracts and not the design of the algorithm.

We find the detections lead major changes in the Estimate at Complete (EAC) by as much as twelve months. The percentage of detections for major EAC changes increases as the time between detection and EAC decreases.

Lastly, approximately 77% of simultaneous changes detected for the CPI and SPI corresponded to large EAC increases. Sequential CPI-SPI detections did not yield any major future EAC changes.

One noteworthy issue we encountered during this analysis was what actually constitutes a problem in contract performance. We used EAC as a problem confirmation measure, but EAC as a problem indicator presented difficulty. The difficulty was EACs may increase because contracts run over cost or because the contract took on a larger

scope and requirements. We differentiated between overrun increases from scope increases by categorizing EAC growth percentages given detection or no detection. If the algorithm did not detect a change in the CPI or SPI and a large percentage increase in EAC occurred, we assumed the increase in EAC was scope-related.

Policy Implications

The ability to detect problems in acquisition contracts offers DoD leadership a method to monitor cost and schedule performance in real-time. The benefit of real-time analysis in defense acquisition is two-fold. First, the identification of contracts which transform suddenly—and significantly-- from good or normal performance to bad performance offers a great capability to program managers and DoD leadership. With real-time problem information, these leaders can identify, isolate, and potentially avoid major cost and schedule overruns. In the future, major cost and schedule overruns may pose serious concerns for acquisition contracts due to the likelihood of greater fiscal scrutiny.

Second, automated real-time analysis helps solve a principal concern of many acquisition leaders. Specifically, automated analysis alleviates some of the strains caused by low personnel levels in the acquisition workforce. To be clear, this does not remove the responsibility of potential users to understand the limitations of this algorithm and method. The algorithm and method provide a way to gain insight into an acquisition contract in addition to or in absence of other information and acquisition professionals.

Follow-On Research

We applied grounded mathematical techniques to a new area of research and data. We used the information available readily in WBS level 1 to reduce collection time. As DCARC or other databases are populated with more contracts that have lower WBS levels, the algorithm and general methodology proposed in this study may find results with more accurate detections and detection lead times. Specifically, two lower level WBS elements common to seemingly all RDT&E contracts are variations of “Design” and “Test”. We began analysis on these WBS elements in our research database, but time constrained our ability to conduct a full analysis. Intuitively, changes in Design and Test affect the overall progress of the program significantly.

Another area of future research follows directly from the results of this analysis. The sensitivity of the detection algorithm should be tied to the tradeoff between 1) the savings from successful detection and overrun mitigation and 2) the cost of potential detection protocol. That is, if a change is detected, what procedures are used to investigate the detection, and at what cost?

Appendix A: Descriptions and Equations of Earned Value (EV) Metrics

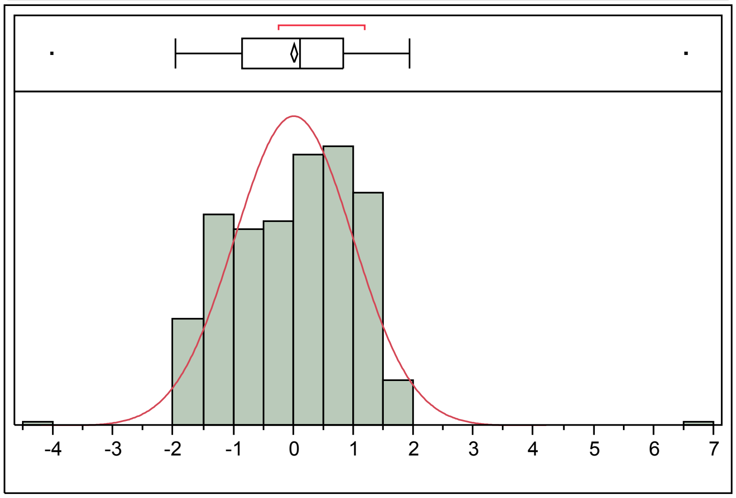
EVM Measure	Description
Actual Cost of Work Performed (ACWP)	Cost of work accomplished
Budgeted Cost of Work Performed (BCWP)	Value of work accomplished
Budgeted Cost of Work Scheduled (BCWS)	Value of work planned
Budget At Completion (BAC)	Total budget for entire contract
Estimate At Completion (EAC)	Estimate of total cost for entire contract
Performance Measurement Baseline (PMB)	Contract time-phased budget plan
Latest Revised Estimate (LRE)	An EAC

Descriptive EVM Measures	Equation	Interpretation
Cost Variance (CV$)	$CV\$ = BCWP - ACWP$	Difference between value and cost of work accomplished
Normalized Cost Variance (NCV)	$NCV = \frac{CV\$}{BAC}$	Cost Variance relative to contract size
Percent Cost Variance (CV%)	$CV\% = \frac{CV\$}{BCWP} * 100$	Shows over and under budget
Schedule Variance (SV$)	$SV\$ = BCWP - BCWS$	Difference between value of work accomplished and value scheduled
Schedule Variance (SVMonths)	$SVMonths = \frac{SV\$}{BCWS}$	Provides a time value for work finished ahead and behind schedule
Normalized Schedule Variance (NSV)	$NSV = \frac{SV\$}{BAC}$	Schedule Variance relative to contract size
Percent Schedule Variance (SV%)	$SV\% = \frac{SV\$}{BCWS} * 100$	Shows ahead and behind schedule
Variance At Completion (VAC)	$VAC = BAC - EAC$	Difference between cost budgeted and cost estimated
Cost Performance Index (CPI)	$CPI = \frac{BCWP}{ACWP}$	Compares the budget to the amount of money spent
Schedule Performance Index (SPI)	$SPI = \frac{BCWP}{BCWS}$	Compares actual value to the value plan
Schedule Cost Index (SCI)	$SCI = CPI * SPI$	
Composite Index (CMI)	$CMI = \alpha CPI + \beta SPI$	
To Complete Performance Index ($TCPI_{EAC}$)	$TCPI = \frac{(BAC - BCWP_{CUM})}{(EAC - ACWP_{CUM})}$	Measures cost efficiency requirement to complete on-budget
Percent Complete (BAC)	$\%Complete = \left(\frac{BCWP_{CUM}}{BAC}\right) * 100$	Compares work plan to program budget
Percent Complete (Months)	$\%Complete = \left(\frac{Months\ from\ Start\ Date}{Total\ Months\ of\ Contract}\right) * 100$	Compares the amount of time spent for a contract to the total amount of time

Appendix B: Distributions of Standardized Residuals ACWP, BCWP, and BCWS

Distributions

standardized residuals_all_acwp



Normal(-4e-17,0.98575)

Quantiles

100.0%	maximum	6.53162
99.5%		1.79547
97.5%		1.53037
90.0%		1.22092
75.0%	quartile	0.82442
50.0%	median	0.10339
25.0%	quartile	-0.8482
10.0%		-1.3658
2.5%		-1.7373
0.5%		-1.8887
0.0%	minimum	-4.042

Moments

Mean	-4.39e-17
Std Dev	0.9857475
Std Err Mean	0.0302627
Upper 95% Mean	0.0593816
Lower 95% Mean	-0.059382
N	1061

Fitted Normal

Parameter Estimates

Type	Parameter	Estimate	Lower 95%	Upper 95%
Location	μ	-4.39e-17	-0.059382	0.0593816
Dispersion	σ	0.9857475	0.9455168	1.0295803

-2log(Likelihood) = 2979.52614511906

Goodness-of-Fit Test

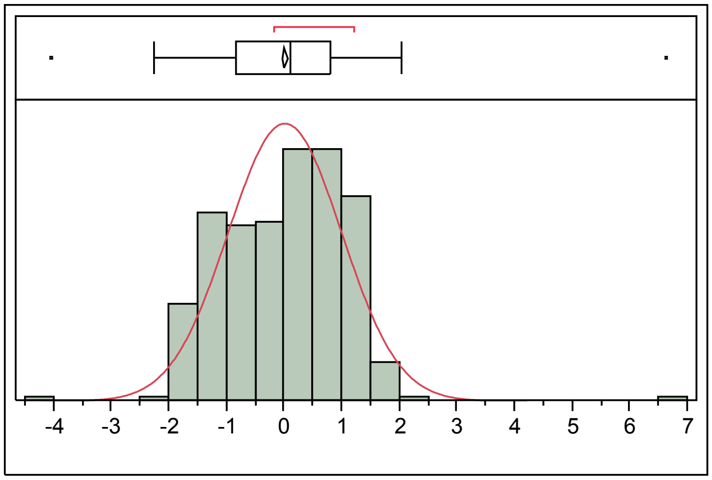
Shapiro-Wilk W Test

W	Prob<W
0.958503	<.0001*

Note: Ho = The data is from the Normal distribution. Small p-values reject Ho.

Distributions

standardized residuals_all_bcwp



Normal(1.6e-16,0.98575)

Quantiles

100.0%	maximum	6.61071
99.5%		1.74846
97.5%		1.55043
90.0%		1.21282
75.0%	quartile	0.81226
50.0%	median	0.10534
25.0%	quartile	-0.8356
10.0%		-1.3545
2.5%		-1.7353
0.5%		-1.9783
0.0%	minimum	-4.0468

Moments

Mean	1.622e-16
Std Dev	0.9857475
Std Err Mean	0.0302627
Upper 95% Mean	0.0593816
Lower 95% Mean	-0.059382
N	1061

Fitted Normal

Parameter Estimates

Type	Parameter	Estimate	Lower 95%	Upper 95%
Location	μ	1.622e-16	-0.059382	0.0593816
Dispersion	σ	0.9857475	0.9455168	1.0295803

-2log(Likelihood) = 2979.52614511906

Goodness-of-Fit Test

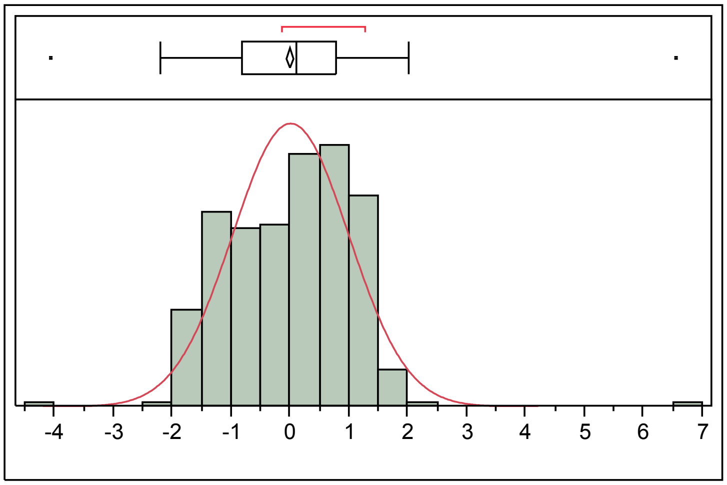
Shapiro-Wilk W Test

W	Prob<W
0.959100	<.0001*

Note: Ho = The data is from the Normal distribution. Small p-values reject Ho.

Distributions

standardized residuals_all_bcws



Normal(2.8e-17,0.98575)

Quantiles

100.0%	maximum	6.53827
99.5%		1.78174
97.5%		1.5131
90.0%		1.21835
75.0%	quartile	0.8028
50.0%	median	0.12092
25.0%	quartile	-0.8239
10.0%		-1.3531
2.5%		-1.7498
0.5%		-1.9559
0.0%	minimum	-4.0574

Moments

Mean	2.846e-17
Std Dev	0.9857475
Std Err Mean	0.0302627
Upper 95% Mean	0.0593816
Lower 95% Mean	-0.059382
N	1061

Fitted Normal

Parameter Estimates

Type	Parameter	Estimate	Lower 95%	Upper 95%
Location	μ	2.846e-17	-0.059382	0.0593816
Dispersion	σ	0.9857475	0.9455168	1.0295803

-2log(Likelihood) = 2979.52614511906

Goodness-of-Fit Test

Shapiro-Wilk W Test

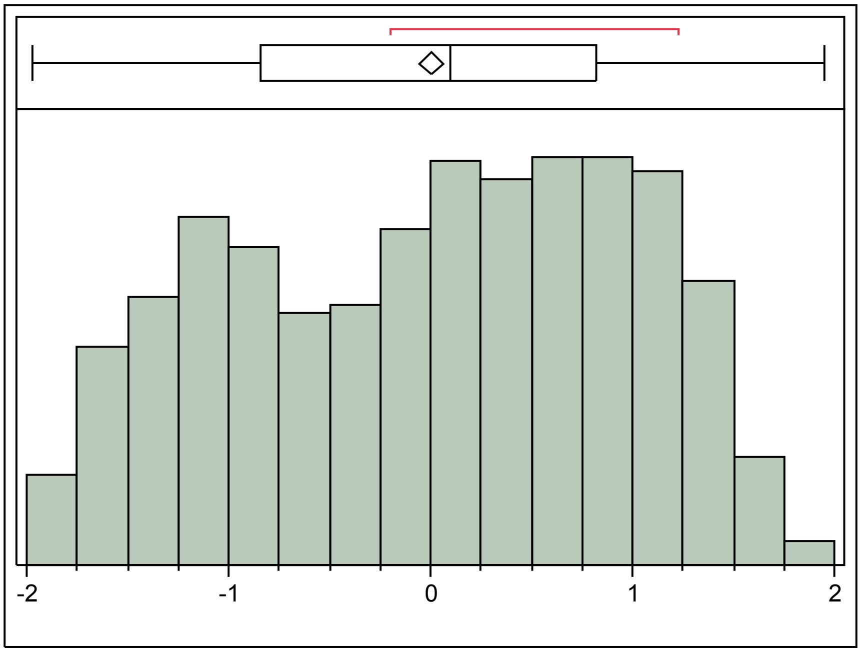
W	Prob<W
0.959594	<.0001*

Note: Ho = The data is from the Normal distribution. Small p-values reject Ho.

Appendix C: Distribution of Standardized Residuals Excluding Statistical Outliers

Distributions

standardized residuals_all_acwp



Quantiles

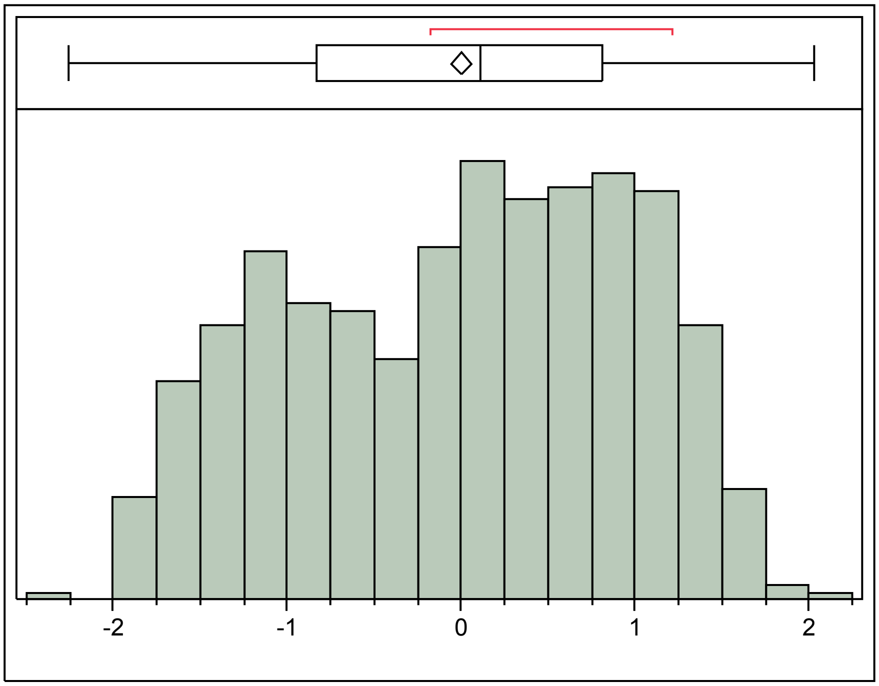
100.0%	maximum	1.95272
99.5%		1.74286
97.5%		1.52856
90.0%		1.22081
75.0%	quartile	0.82369
50.0%	median	0.10339
25.0%	quartile	-0.8462
10.0%		-1.3655
2.5%		-1.7328
0.5%		-1.8802
0.0%	minimum	-1.9699

Moments

Mean	-0.002351
Std Dev	0.9580001
Std Err Mean	0.0294386
Upper 95% Mean	0.0554138
Lower 95% Mean	-0.060116
N	1059

Distributions

standardized residuals_all_bcwp



Quantiles

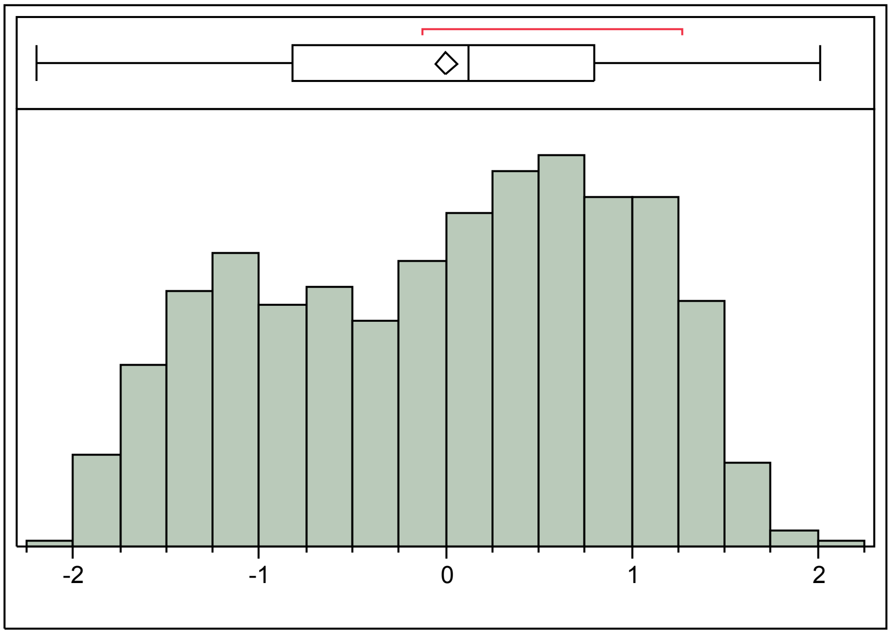
100.0%	maximum	2.0317
99.5%		1.70982
97.5%		1.54031
90.0%		1.21265
75.0%	quartile	0.81179
50.0%	median	0.10534
25.0%	quartile	-0.8332
10.0%		-1.3505
2.5%		-1.7325
0.5%		-1.9444
0.0%	minimum	-2.264

Moments

Mean	-0.002421
Std Dev	0.9574677
Std Err Mean	0.0294223
Upper 95% Mean	0.0553116
Lower 95% Mean	-0.060154
N	1059

Distributions

standardized residuals_all_bcws



Quantiles

100.0%	maximum	2.01195
99.5%		1.74835
97.5%		1.50934
90.0%		1.21417
75.0%	quartile	0.80002
50.0%	median	0.12092
25.0%	quartile	-0.8236
10.0%		-1.3506
2.5%		-1.7421
0.5%		-1.9227
0.0%	minimum	-2.1969

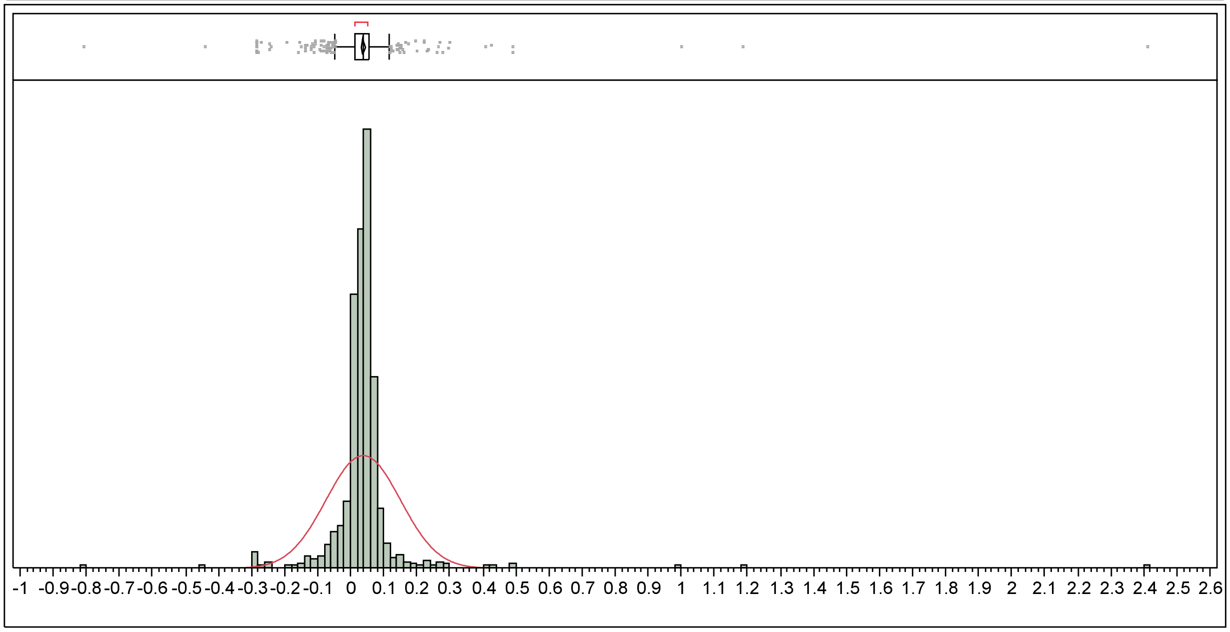
Moments

Mean	-0.002343
Std Dev	0.9578957
Std Err Mean	0.0294354
Upper 95% Mean	0.0554158
Lower 95% Mean	-0.060101
N	1059

Appendix D: Distribution of Standardized Residuals CPI and SPI

Distributions

Std Residual Column 1



Normal(0.03578,0.1126)

Quantiles

100.0%	maximum	2.40862
99.5%		0.47246
97.5%		0.15034
90.0%		0.07656
75.0%	quartile	0.05476
50.0%	median	0.03812
25.0%	quartile	0.01405
10.0%		-0.0155
2.5%		-0.1231
0.5%		-0.2861
0.0%	minimum	-0.809

Moments

Mean	0.0357803
Std Dev	0.1126046
Std Err Mean	0.0034751
Upper 95% Mean	0.0425992
Lower 95% Mean	0.0289615
N	1050

Fitted Normal

Parameter Estimates

Type	Parameter	Estimate	Lower 95%	Upper 95%
Location	μ	0.0357803	0.0289615	0.0425992
Dispersion	σ	0.1126046	0.1079859	0.1176391

-2log(Likelihood) = -1607.36141473001

Goodness-of-Fit Test

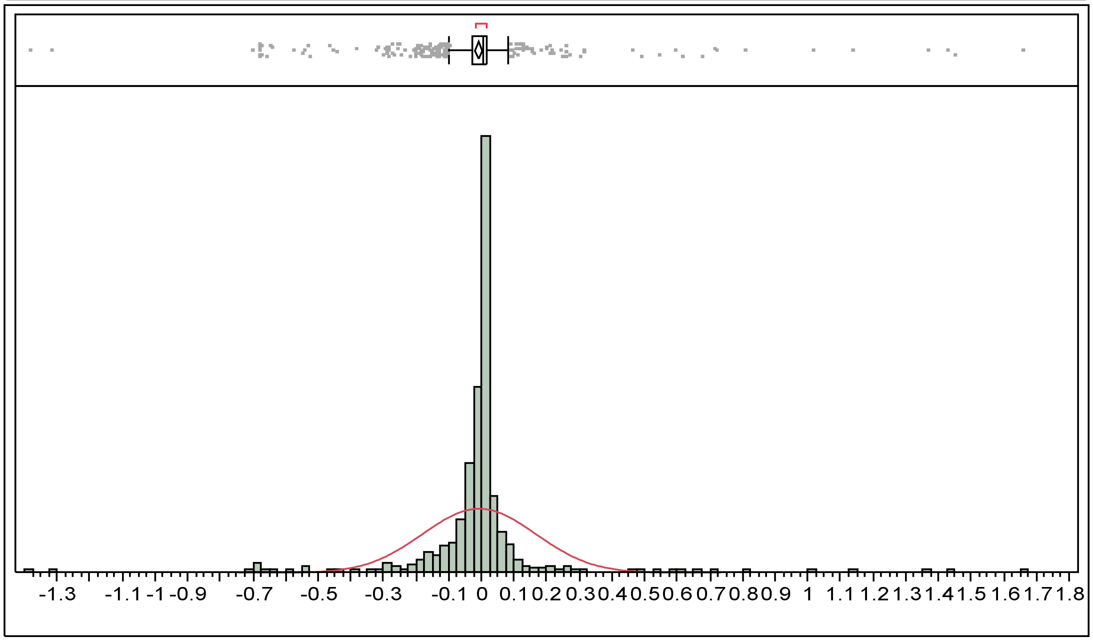
Shapiro-Wilk W Test

W	Prob<W
0.415546	<.0001*

Note: Ho = The data is from the Normal distribution. Small p-values reject Ho.

Distributions

Std Residual Column 2



Quantiles

100.0%	maximum	1.65988
99.5%		1.10807
97.5%		0.21802
90.0%		0.05448
75.0%	quartile	0.01618
50.0%	median	0.00362
25.0%	quartile	-0.0297
10.0%		-0.109
2.5%		-0.301
0.5%		-0.6789
0.0%	minimum	-1.3822

Moments

Mean	-0.009432
Std Dev	0.1734183
Std Err Mean	0.0053646
Upper 95% Mean	0.0010948
Lower 95% Mean	-0.019958
N	1045

Fitted Normal

Parameter Estimates

Type	Parameter	Estimate	Lower 95%	Upper 95%
Location	μ	-0.009432	-0.019958	0.0010948
Dispersion	σ	0.1734183	0.1662889	0.1811911

-2log(Likelihood) = -697.200685238011

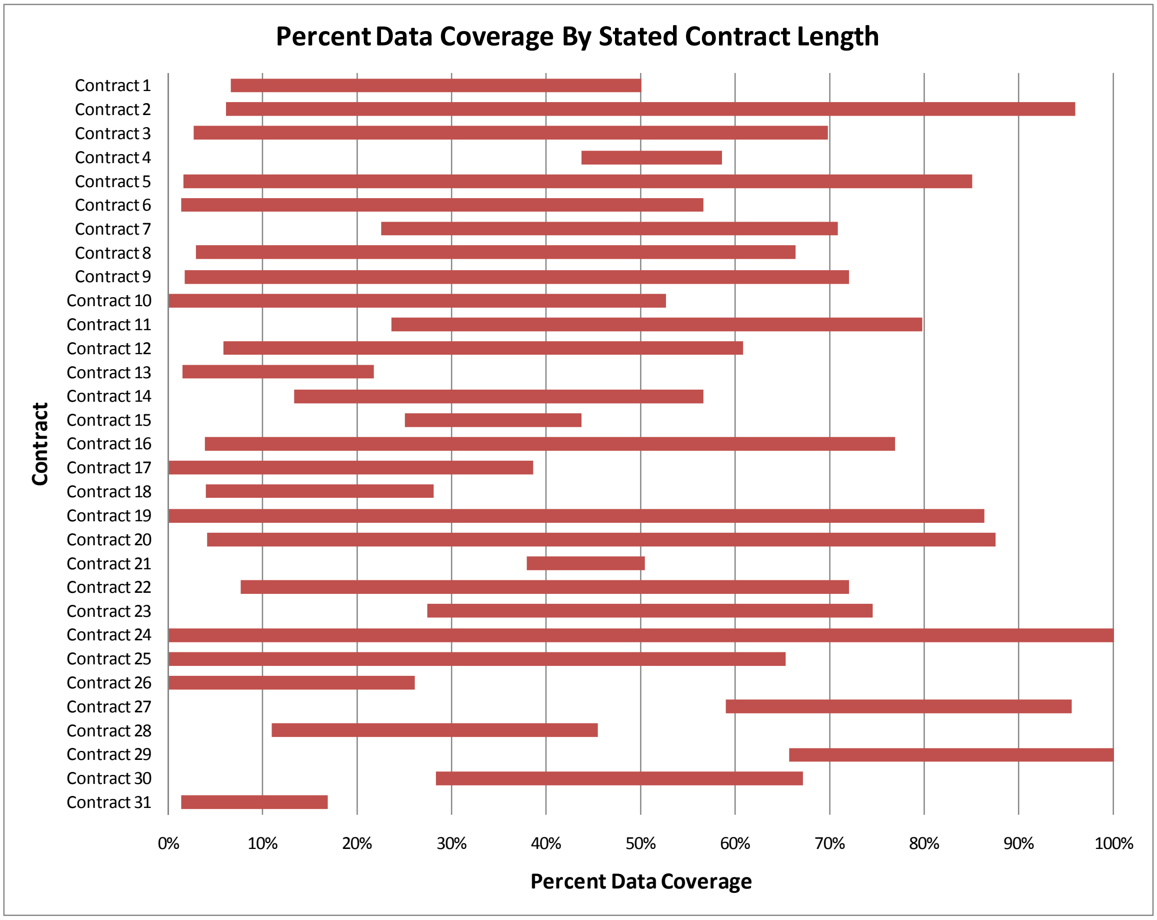
Goodness-of-Fit Test

Shapiro-Wilk W Test

W	Prob<W
0.537532	<.0001*

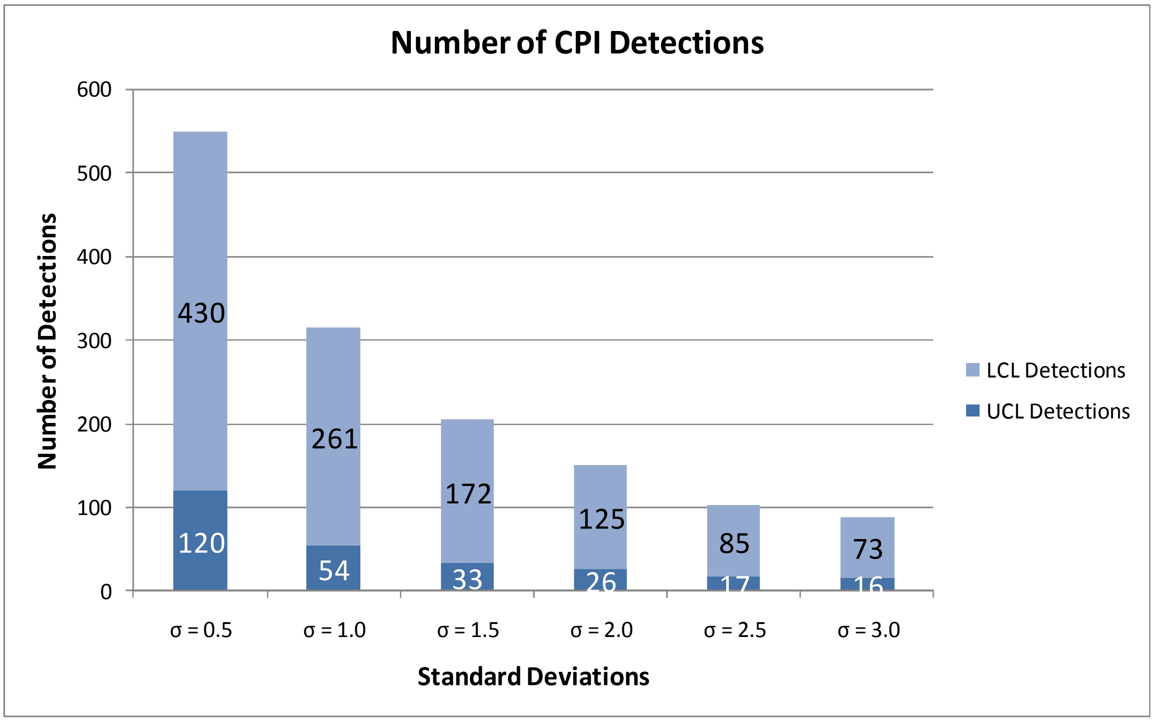
Note: Ho = The data is from the Normal distribution. Small p-values reject Ho.

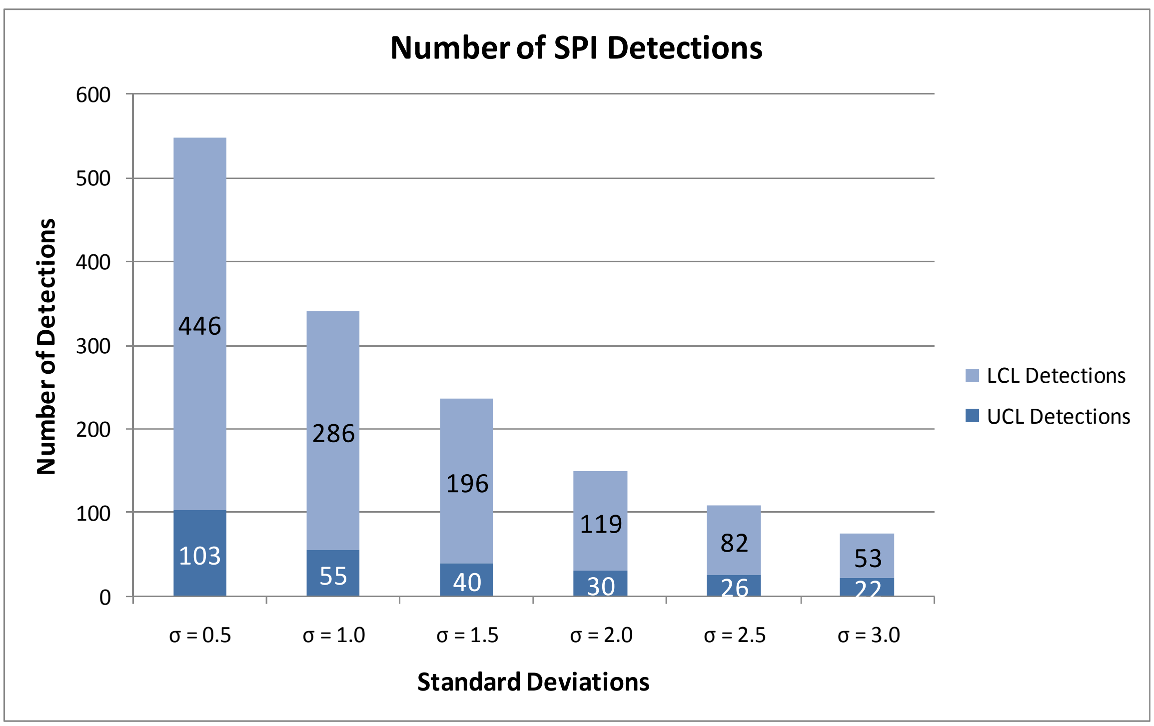
Appendix E: Data Coverage

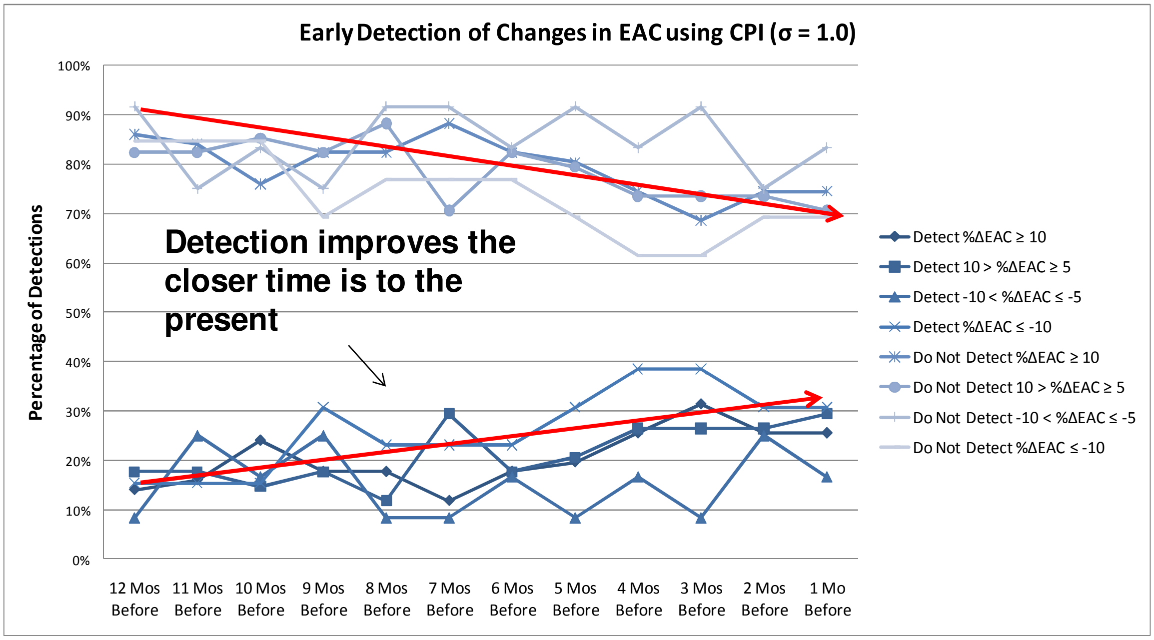


[adapted from (Rosado, 2011)]

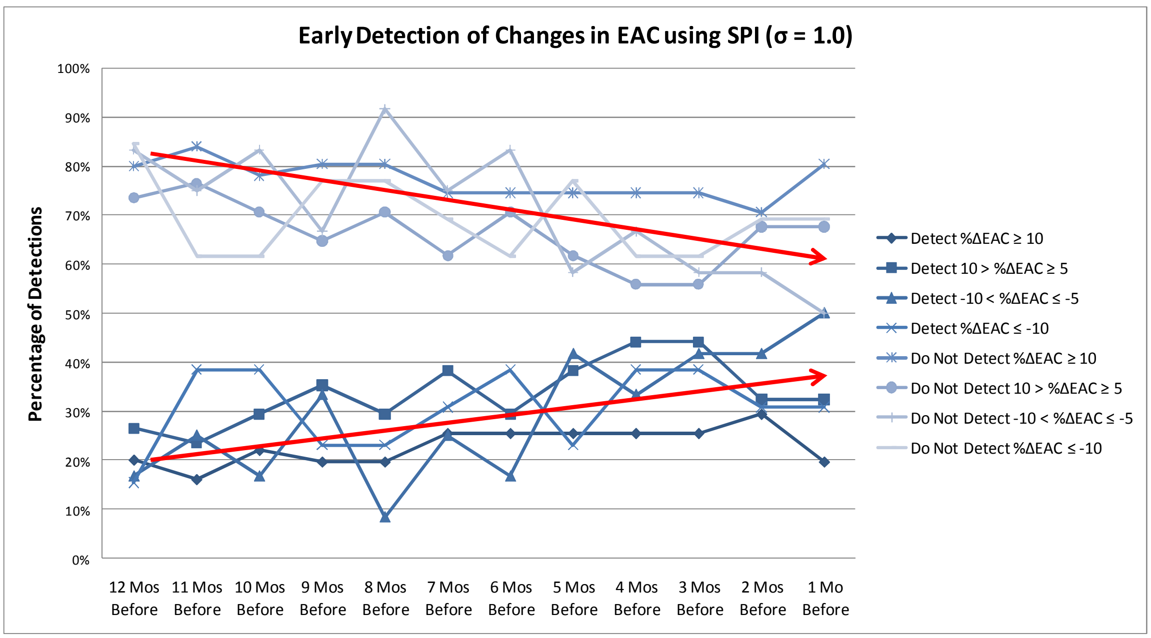
Appendix F: Change Detection Results







Early Detection of Changes in EAC using CPI (σ = 1.0)
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
Percentage of Detections
Detection improves the closer time is to the present
12 Mos Before
11 Mos Before
10 Mos Before
9 Mos Before
8 Mos Before
7 Mos Before
6 Mos Before
5 Mos Before
4 Mos Before
3 Mos Before
2 Mos Before
1 Mo Before
Detect %ΔEAC ≥ 10
Detect 10 > %ΔEAC ≥ 5
Detect -10 < %ΔEAC ≤ -5
Detect %ΔEAC ≤ -10
Do Not Detect %ΔEAC ≥ 10
Do Not Detect 10 > %ΔEAC ≥ 5
Do Not Detect -10 < %ΔEAC ≤ -5
Do Not Detect %ΔEAC ≤ -10



Early Detection of Changes in EAC using SPI (σ = 1.0)
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
Percentage of Detections
12 Mos Before
11 Mos Before
10 Mos Before
9 Mos Before
8 Mos Before
7 Mos Before
6 Mos Before
5 Mos Before
4 Mos Before
3 Mos Before
2 Mos Before
1 Mo Before
Detect %ΔEAC ≥ 10
Detect 10 > %ΔEAC ≥ 5
Detect -10 < %ΔEAC ≤ -5
Detect %ΔEAC ≤ -10
Do Not Detect %ΔEAC ≥ 10
Do Not Detect 10 > %ΔEAC ≥ 5
Do Not Detect -10 < %ΔEAC ≤ -5
Do Not Detect %ΔEAC ≤ -10

Months Before	# CPI Leads SPI	# SPI Leads CPI
1	174	161
2	158	149
3	147	140
4	125	130
5	118	121
6	106	111
7	97	116
8	93	110
9	87	107
10	87	106
11	82	104
12	83	92

Bibliography

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control , 19*.

Basseville, M., & Nikiforov, I. V. (1993). *Detection of Abrupt Changes: Theory and Application.* Englewood Cliffs NJ: PTR Prentice Hall, Inc.

Borodin, A., & El-Yaniv, R. (1998). *Online Computation and Competitive Anlaysis.* Cambridge University Press.

Borodkin, L., & Mottl', V. (1976). Algorithm for Finding the Jump Times of Random Process Equation Parameters. *Automation and Remote Control , 37* (6), 23-32.

Box, G. E., & Pierce, D. A. (1970). Distribution of the residual autocorrelations in autoregressive-integrated moving-average time series models. *Journal of the American Statistical Association , 65*, 1509-1526.

Box, G. E., Jenkins, G. M., & Reinsel, G. C. (1994). *Time Series Analysis: Forecasting and Control.* Upper Saddle River NJ: Prentice Hall.

Cios, K. J., & Moore, G. W. (2002). Uniqueness of Medical Data Mining. *Artificial Intelligence in Medicine , 26* (1), 1-24.

Cohen, A. (1987). *Biomedical Signal Processing* (Vol. 1 and 2). Boca Raton FL: CRC Press.

Dasgupta, D., & Forrest, S. (1996). Novelty Detection in Time Series data using Ideas from Immunology. *5th International Conference on Intelligent Systems.* Reno NV.

Duncan, A. (1986). *Quality Control and Industrial Statistics.* Homewood IL: Richard D. Irwin, Inc.

Gustafsson, F. (2000). *Adaptive Filtering and Change Detection.* West Sussex England: John Wiley & Sons Ltd.

Headquarters Air Force Materiel Command, Financial Management. (1994). *Guide to Analysis of Contractor Cost Data.*

Ljung, G., & Box, G. (1978). On a measure of lack of fit in time series models. *Biometrika* (65), 297-303.

Makridakis, S., Wheelwright, S. C., & Hyndman, R. J. (1998). *Forecasting: Methods and Applications.* New York NY: John Wiley & Sons Inc.

Microsoft Corporation. (2011, January 23). Excel 2007 Help and How-to.

Morin, J. M. (2010). Achieving Acquisition Excellence in the Air Force: A Financial Management Perspective. *Armed Forces Comptroller , 55* (2).

Nau, R. F. (2005, May 16). *Forecasting*. Retrieved January 16, 2011, from Decision 411: http://www.duke.edu/~rnau/411home.htm

NDIA/PMSC. (2009). *Earned Value Management Systems Intent Guide.*

Newbold, P., Carlson, W. L., & Thorne, B. (2010). *Statistics for Business and Economics.* Upper Saddle River NJ: Prentice Hall.

Nisevich, S. (2010, July 20). B-2 Earned Value Analyst. (G. Keaton, Interviewer)

OUSD(AT&L)ARA/AM(SO). (2005). *Data Item Description- Contract Performance Report (CPR).* OUSD(AT&L).

Rosado, W. (2011). *Masters Thesis* . Wright-Patterson AFB OH, United States of America: Air Force Institute of Technology.

SAF/FMC. (2010, January 8). USAF Inflation Indices Base Year 2010.

Schwarz, G. E. (1978). Estimating the dimension of a model. *Annals of Statistics , 6* (2), 461-464.

Shapiro, S., & Wilk, M. (1965). An analysis of variance test for normaility (complete samples). *Biometrika* (52), 591-611.

Shewhart, W. (1931). *Economic Control of Quality Manufactured Product.* Princeton NJ: D. Van Nostrand.

Venkatesh, S. (2007). Recognition of Emergent Human Behaviour in a Smart HomeL A Data Mining Approach. *Pervasive and Mobile Computing , 3* (2), 95-116.

Willsky, A. S. (1976). A Survey of Design Methods for Failure Detection in Dynamic Systems. *Automatica , 12*, 601-611.

Woodward, R., & Goldsmith, P. (1964). *Cumulative Sum Techniques.* Edinburgh UK: Oliver and Boyd.

REPORT DOCUMENTATION PAGE			*Form Approved OMB No. 074-0188*
The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of the collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to an penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**			
1. REPORT DATE *(DD-MM-YYYY)* 24-03-2011	**2. REPORT TYPE** Master's Thesis		**3. DATES COVERED** *(From – To)* Aug 2009-Mar 2011
4. TITLE AND SUBTITLE Using Earned Value Data to Detect Potential Problems in Acquisition Contracts			**5a. CONTRACT NUMBER**
			5b. GRANT NUMBER
			5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) Keaton, C. Grant, 2d Lt, USAF			**5d. PROJECT NUMBER** JON 11-269
			5e. TASK NUMBER
			5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S) Air Force Institute of Technology Graduate School of Engineering and Management (AFIT/EN) 2950 Hobson Way WPAFB OH 45433-7765			**8. PERFORMING ORGANIZATION REPORT NUMBER** AFIT/GCA/ENC/11-01
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Mr. Steve Miller Office of the Secretary of Defense, Cost Assessment and Program Evaluation Room BE-829 1800 Defense Pentagon Washington, DC 20301-1800 (703) 697-5056			**10. SPONSOR/MONITOR'S ACRONYM(S)** OSD/CAPE
			11. SPONSOR/MONITOR'S REPORT NUMBER(S)
12. DISTRIBUTION/AVAILABILITY STATEMENT APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED			
13. SUPPLEMENTARY NOTES This material is declared a work of the United States Government and is not subject to copyright protection in the United States.			
14. ABSTRACT Government contractors report earned value (EV) information to government agencies in monthly Contract Performance Reports (CPR). Though major differences may exist in the data between subsequent CPRs, we know of no government effort to detect these occurrences. The identification of major changes may locate and isolate problems and thus prevent million and billion dollar cost and schedule overruns. In this study, we develop an approach to identify changes in the Cost Performance Index (CPI) and the Schedule Performance Index (SPI) that may indicate problems with contract performance. We find the detection algorithm indentifies changes in the CPI and the SPI that correspond to large future changes in the Estimate at Complete (EAC). The ability to detect unusual changes provides decision-makers with warnings for potential problems for acquisition contracts.			
15. SUBJECT TERMS problem detection, acquisition contracts, earned value management (EVM)			

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON** Dr. Edward White (AFIT/ENC)
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER** *(Include area code)*
U	U	U	UU	77	937-255-3636 ext 4540, Edward.white@afit.edu

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39-18

	Form Approved OMB No. 074-0188